Article

Convergence of evolving artificial intelligence and machine learning techniques in precision oncology

Check for updates

Elena Fountzilas [1,4], Tillman Pearce[2,4], Mehmet A. Baysal [3], Abhijit Chakraborty [3] & Apostolia M. Tsimberidou [3] ✉

The confluence of new technologies with artificial intelligence (AI) and machine learning (ML) analytical techniques is rapidly advancing the field of precision oncology, promising to improve diagnostic approaches and therapeutic strategies for patients with cancer. By analyzing multi-dimensional, multiomic, spatial pathology, and radiomic data, these technologies enable a deeper understanding of the intricate molecular pathways, aiding in the identification of critical nodes within the tumor's biology to optimize treatment selection. The applications of AI/ML in precision oncology are extensive and include the generation of synthetic data, e.g., digital twins, in order to provide the necessary information to design or expedite the conduct of clinical trials. Currently, many operational and technical challenges exist related to data technology, engineering, and storage; algorithm development and structures; quality and quantity of the data and the analytical pipeline; data sharing and generalizability; and the incorporation of these technologies into the current clinical workflow and reimbursement models.

Artificial intelligence (AI) refers to the ability of a machine or computational model to recognize or "learn" patterns and relationships from input of representative examples (training data) and make accurate predictions regarding independent, or previously unseen, data[1].

Early efforts at designing AI systems, considered "symbolic" or "rules-based" AI, were based on encoding human knowledge into computer programs, so-called "expert systems," that worked for "narrow tasks" but not for complex tasks. An example is IBM Watson for Oncology, an AI system for clinical decision-making in oncology, that failed to achieve high concordance with expert clinicians in terms of treatment recommendations[2–4].

Machine learning (ML) emphasizes the process by which a computer system or model "learns" or improves its predictive performance by discovering patterns in data and incorporating what has been learned into the model in an iterative fashion[5]. ML techniques are considered "supervised learning," with predetermined outputs, or unsupervised learning (e.g., without the need for explicit labeling or prior knowledge of the data), to discover underlying data patterns that are not evident to humans[6].

Deep learning (DL) is a subset of ML that focuses on artificial neural networks, such as convolutional neural networks (CNNs). DL models have improved the state of the art in various fields, including computer vision

(e.g., pathology and radiology image analysis), natural language processing (NLP) (e.g., electronic health record [EHR] mining), and speech recognition. DL has been used for facial recognition, image classification, and video, speech, and audio processing[1,7].

Foundation models or large language models (LLMs) such as generative pre-trained transformers (GPTs) and vision transformers, the most recent fundamental advances in DL-based NLP, were first described in 2017[8,9]. LLMs can enable humans to interact directly with a computer using natural language (e.g., English). Foundation models are "pretrained" on vast amounts of data from disparate sources, such as internet-derived digital data. The models learn to identify objects from the input data, and through "transfer learning," their capacity to recognize objects can be fine-tuned for specific downstream tasks, such as recognizing cancer cells from a whole slide image of a tumor biopsy. Foundation models have the capacity for "self-supervised" learning, e.g., the pre-training task is derived automatically from unannotated or unlabeled data, a promising feature for the analysis of oncology datasets. Importantly, foundation models can accommodate multiple types, or "modes," of data (e.g., text, imaging, pathology, molecular biology, video, audio), incorporating them into a prediction and enabling "multimodal" analysis that has potential applications for

[1]Department of Medical Oncology, St Luke's Clinic, Panorama, Thessaloniki, Greece. [2]TCellCo, 415 De Haro Street, San Francisco, CA, USA. [3]Department of Investigational Cancer Therapeutics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX, USA. [4]These authors contributed equally: Elena Fountzilas, Tillman Pearce. ✉e-mail: atsimber@mdanderson.org

decision-making in oncology[10]. This is particularly important for measuring biological markers and disease.

These AI approaches are both distinct from and complementary to traditional inferential statistics. Both inferential statistics and AI approaches can advance precision oncology, which refers to the use of information about a patient's genes, proteins, and environment to diagnose and treat disease. Initially, the term "precision oncology" was used to describe targeting tumor molecular abnormalities with drugs known to inhibit the function of a molecular alteration. In recent years, precision oncology has included the development of therapeutic agents that target any biological abnormality that is associated with carcinogenesis. Consequently, owing to recent major breakthroughs in immunotherapeutic strategies, the armamentarium of the precision medicine approach now also includes immunotherapy. The term immuno-oncology refers to the use of immunotherapeutic approaches that include immune checkpoint inhibitors, chimeric antigen receptor T-cell (CAR-T) therapy, cytokines, and vaccines to treat patients with cancer[11]. Immuno-oncology is used without stringent biomarker selection, in contrast to the use of targeted therapies with small molecules. In a recent meta-analysis, the use of checkpoint inhibitors was associated with higher rates of overall response, progression-free survival (PFS), and overall survival (OS) in patients with biomarker-positive tumors compared with those with biomarker-negative tumors[12]. The efficacy of immunotherapeutic treatments varies among different patients and tumor types, underlying the importance of exploring the complex immune system in each patient, discovering potential mechanisms of response and resistance to these therapeutic approaches, and identifying predictive biomarkers that will enable the selection of the optimal immunotherapy approach for each patient. The goal of precision immuno-oncology is the optimization of cancer immunotherapy based on the individual characteristics of each patient, in combination with specific genetic, molecular, and immunological characteristics of the patient's tumor, to increase efficacy while minimizing toxicity. The application of AI/ML in precision oncology may enable the analysis of big "omics" data in combination with clinical, pathological, treatment, and outcome data, providing sophisticated and powerful tools to optimize the development of biomarkers and treatment of patients.

New modalities for deep measurement of disease (multiplex digital spatial analysis of pathology slides, quantitative digital analysis of medical images, genomic sequencing, and mass spectrometric analysis of biologic molecules) create analytical challenges due to the high-dimensional, multimodal nature of the data. To accelerate the use of these analytical tools in precision oncology, the key task for oncologists is to ensure that the tools are adapted to the intended goals and available data (Fig. 1). AI/ML is applied in many areas of oncology, including generative work, NLP, and other structured sources of data like the EHR. In this clinically focused overview, we provide a technological and clinical perspective on the use of AI/ML in precision oncology to increase our understanding of tumor biology and to aid in the development of biomarkers that improve treatment selection in patients with cancer.

Methodology

We conducted a PubMed search using the terms "artificial intelligence" and "precision oncology" and another search using the terms "machine learning" and "precision oncology" and publication dates from January 1, 2020, through November 30, 2024. The following filters were used: "Clinical Study"; "Clinical Trial"; and "Clinical Trial Phase I." Phase II or Phase III clinical trials were included in the term "Clinical Trial." Clinical trials that were not cancer-related ($n = 6$) or did not include AI-based analyses ($n = 3$) or were not original studies ($n = 1$) were excluded. The time period was selected before data extraction began.

In addition to the above search, we have reviewed published articles that utilized AI/ML methodologies to analyze patient-derived data, using the following criteria: detailed description of AI/ML methodology, inclusion of patient data, and results that provided potentially novel clinical insights not achieved by conventional methodologies.

Results

Using the criteria listed in the Methodology section, 20 trials utilized AI/ML methodologies to analyze patient-derived data across diverse tumor types (Table 1). These 20 studies aimed to identify models that improved diagnostic accuracy[13–17], improved the prediction of clinical outcomes[18–26], explored the tumor molecular profile[27,28] or were related to patient care[29–32]. Common limitations included retrospective study design, small sample size, and lack of external validation. Overall, these studies exemplify the transformative effect of AI/ML tools on the diagnosis, treatment, and individualized management of cancer, with the hope of optimizing patient care.

Application of AI/ML in Precision Oncology
Digital pathology

There are multiple areas within the field of digital pathology where AI/ML is being explored. Key applications include automation in immunohistochemistry (IHC) scoring, the inference of clinically relevant features beyond histology from hematoxylin and eosin (H&E) images, and novel insights from emerging tools for measuring multiplex, single-cell, and spatially resolved analytes from tumor tissue.
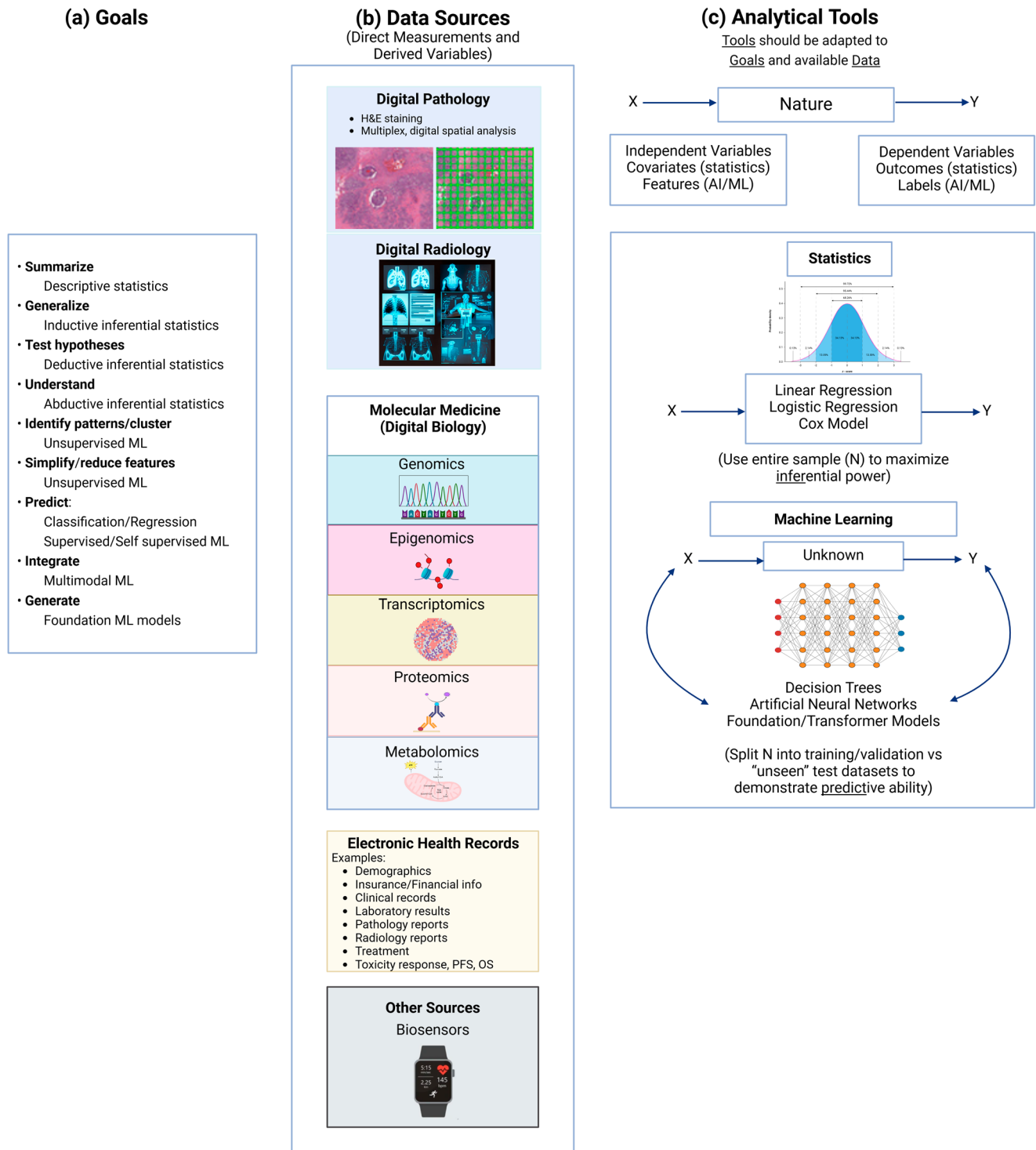
The role of AI in automating IHC biomarker scoring. AI-based technology may help standardize IHC assessments, including those used in routine practice for treatment selection based on biomarkers (e.g., PD-L1, HER2, ER, PR, Ki-67). This would be especially valuable as an assistance tool for pathologists because the standard manual approach is time consuming and is associated with high intra-observer variability[33–35]. An automated and quantitative AI-based technology has the potential to standardize the quality of patient care across centers and geographic areas by overcoming variability in assessment by pathologists, specifically in rare and complex cases, increasing accuracy and reproducibility, and reducing turnaround time[33,36–39].

Automated AI-based IHC scoring systems have been evaluated by analyzing scans of whole-slide images (WSIs) of tumor samples in settings where the standard of care currently requires manual determination of protein expression by IHC[33,40–45]. For example, several independent groups have demonstrated the potential of AI-supported quantitative PD-L1 evaluation using CNNs[38,40,41]. Two separate groups developed CNN systems that were able to automatically detect the tumor area within WSIs and to calculate the IHC-based PD-L1 tumor proportion score (TPS) with high consistency between the AI systems and pathologists[40,41].

Others developed a similar CNN PD-L1 TPS classifier and retrospectively analyzed 1746 samples across CheckMate studies of nivolumab combined with ipilimumab for the treatment of patients with various cancers[38]. The automated AI system classified more patients as PD-L1 positive (at both the 1% and 5% expression levels) compared with manual scoring in most tumor types. Importantly, similar improvements in response and survival were observed using both AI-powered and manual scoring. However, automated AI-powered digital analysis may identify more patients who would benefit from immunotherapy treatment compared with manual assessment[38]. This is because AI-powered methods can analyze larger datasets, detect subtle patterns, and provide more consistent evaluations, potentially reducing the variability inherent in manual assessments.

Recent advances in context-aware attention mechanisms, such as the Context-Aware Multiple Instance Learning (CAMIL) model, have significantly improved diagnostic accuracy in medical imaging. CAMIL prioritizes relevant regions within WSIs by analyzing spatial relationships and contextual interactions between neighboring areas. This approach reduces misclassification rates and enhances diagnostic reliability[46].

The workflow for pathologists in the setting of breast cancer diagnosis is burdensome, as it includes manual quantitative IHC assessment with clinically relevant cutoff levels of multiple proteins including HER2, ER, PR, Ki-67, and PD-L1. HER2 assessment is known to be associated with significant diagnostic variability. Intra-tumoral heterogeneity within WSIs of tumor tissue hinders the accurate identification of all cells expressing the

(a) Goals

(b) Data Sources
(Direct Measurements and Derived Variables)

(c) Analytical Tools
Tools should be adapted to Goals and available Data

- **Summarize**
  Descriptive statistics
- **Generalize**
  Inductive inferential statistics
- **Test hypotheses**
  Deductive inferential statistics
- **Understand**
  Abductive inferential statistics
- **Identify patterns/cluster**
  Unsupervised ML
- **Simplify/reduce features**
  Unsupervised ML
- **Predict**:
  Classification/Regression
  Supervised/Self supervised ML
- **Integrate**
  Multimodal ML
- **Generate**
  Foundation ML models

Digital Pathology
- H&E staining
- Multiplex, digital spatial analysis

Digital Radiology

Molecular Medicine (Digital Biology)

Genomics

Epigenomics

Transcriptomics

Proteomics

Metabolomics

Electronic Health Records
Examples:
- Demographics
- Insurance/Financial info
- Clinical records
- Laboratory results
- Pathology reports
- Radiology reports
- Treatment
- Toxicity response, PFS, OS

Other Sources
Biosensors

X ⟶ Nature ⟶ Y

Independent Variables
Covariates (statistics)
Features (AI/ML)

Dependent Variables
Outcomes (statistics)
Labels (AI/ML)

Statistics

X ⟶ Linear Regression / Logistic Regression / Cox Model ⟶ Y

(Use entire sample (N) to maximize inferential power)

Machine Learning

X ⟶ Unknown ⟶ Y

Decision Trees
Artificial Neural Networks
Foundation/Transformer Models

(Split N into training/validation vs "unseen" test datasets to demonstrate predictive ability)

respective protein. In addition, manual counting of tumor cells to evaluate biomarker expression levels is associated with low efficiency and poor reproducibility. In clinical practice, the training and experience of pathologists significantly influence the accuracy of biomarker assessment (i.e., PD-L1 expression)[47,48]. For instance, untrained pathologists exhibit lower intraclass concordance in PD-L1 expression compared to their highly trained colleagues[41].

One group assessed various ML and DL approaches to automated quantitative HER2 IHC scoring and found that a CNN model outperformed classical ML approaches[49]. Using 71 breast tumor samples, a concordance of 83% between the automated scoring system and a pathologist's assessment

was demonstrated. Discordance between automated and manual scoring was found to be associated with HER2 staining heterogeneity in these cases; notably, an independent review of the discordant cases led to a modification of the initial pathologist assessment in 8/12 cases, highlighting the potential utility of AI assistance for the identification of ambiguous cases[49].

This potential benefit of using AI as an assistance tool was demonstrated in a separate study using a CNN to classify cells as either tumor or non-tumor and to quantify IHC staining intensity for ER/PR and Ki67[42]. The goal of the study was to evaluate the reliability of using an AI system as a diagnostic decision support tool in a routine clinical pathology setting (6 WSI scanners/microscopes; 3 staining machines; manual scoring by 10

Fig. 1 | Clinical perspective on the use of AI/ML in precision oncology. Analytical tools must be adapted to intended goals and available data. Innovations in artificial intelligence, machine learning analytical techniques, and new modalities for deep measurement of disease hold great promise for advancing precision oncology. Central to deriving maximal benefit from these innovations is for researchers and practitioners to clearly articulate (**a**) what goals they are seeking to achieve and (**b**) what sources of data are available for analysis. This will then dictate the choice of (**c**) analytical tools. Inferential statistics is a "data model" approach that seeks to understand or infer the relationships between independent variables (covariates) and dependent variables (outcomes) based on prior assumptions about the data structure. In contrast, machine learning is an "algorithmic model" approach, which makes few assumptions about the data but rather designs algorithms that can input direct measurements or derived variables, transform them through the mathematical workings of the algorithm into "features", and ultimately "learn" to predict the dependent variable (label). Inference is to statistics as prediction is to machine learning, and moving forward, we will need to use all the tools in our analytical toolkit. Interventional statisticians (i.e., clinical trialists) often use the entire sample size for a primary analysis to maximize the power of the analysis and less frequently use training and validation sets, whereas data scientists and observational statisticians (i.e., epidemiologists) divide patient samples into training, validation, and test sets to demonstrate predictive ability on the "unseen" test set based on analysis of the training and validation sets. Both utilize models, but the primary objectives are different[187]. Inferential statistics, using "data models," seeks to understand or infer the relationships between the independent variables and the dependent or outcome variables within a dataset in three fashions: exploratory or inductive, hypothesis-testing or deductive, and explanatory or abductive. In all cases, a model that makes assumptions about the structure of the data (normal distribution or proportional hazards between groups) is applied to the dataset in order to understand the relationship between prespecified independent input variables ("x") and dependent outcome variables ("y") and to draw population inferences from a sample[188]. ML "algorithmic models" often make fewer assumptions compared to inferential statistics about the structure of the data or the nature of the relationship between variables. The flexibility of ML "algorithmic models" lies in their ability to adapt these assumptions based on the chosen model and application, making these models applicable to a wide range of predictive tasks[189]. Since ML/DL is a form of "representation learning," in that the machine is fed raw data and develops its own models for pattern recognition[190], the results can be used to make predictions about independent or "unseen" data. "Created with BioRender.com".

pathologists from 8 different centers) by ensuring that the use of AI did not adversely impact the pathologist assessment. Individual AI analysis results were confirmed by pathologists in 95.8% of the Ki-67 cases and 93.2% of the ER/PR cases, indicating the reliability of IHC scoring with the support of the CNN AI tool. Statistical analysis also demonstrated high interobserver variance between pathologists in conventional IHC quantification, which decreased slightly with AI assistance[42].

These reports indicate that AI can assist pathologists by automating IHC scoring, reducing inter-observer variability (a challenge associated with the determination of clinically relevant expression cutoffs), and shortening the diagnostic workup period. Prospective trials are needed to confirm the clinical validity and utility of these promising technologies.

The use of AI to predict biologic characteristics from H&E-stained WSIs. Given that DL models such as CNNs exhibit "representation learning" and are able to extract "deep" patterns from input data, these DL models have demonstrated the ability to reveal molecular characteristics from H&E-stained WSIs, as histology reflects biology[33,50–53]. These DL models have been associated with difficulties in "explaining" how they developed their predictions and offer the opportunity for identifying human-interpretable features (HIFs) based on cell morphology and histological patterns[54]. Investigators have prioritized identification of HIFs derived from CNN models when analyzing H&E WSIs from patients with cancer to predict molecular phenotypes. HIFs were correlated with established markers of the tumor microenvironment that are predictive of diverse molecular signatures, including expression of immune checkpoint proteins and homologous recombination deficiency, indicating that their application should be further explored[54].

DL analysis of H&E images can predict molecular alterations prior to, and potentially in lieu of, performing IHC or molecular confirmatory testing. HER2 and BRCA expression was predicted from H&E-stained WSIs from patients with breast cancer using a CNN that separately processes H&E-stained slide patches or tiles and outputs an IHC label for the WSI[54]. The study demonstrated 83.3% and 53.8% prediction accuracy for HER2 and BRCA, respectively[54]. Similarly promising early results for BRCA prediction have been reported by others[50]. In addition, CNN-based analyses of H&E-stained WSIs have been used to prioritize patients for microsatellite instability (MSI)/mismatch repair deficiency (dMMR) testing to select patients for treatment with immunotherapy[55,56]. A CNN model to predict MSI was trained using 100 H&E-stained WSIs from patients with colorectal cancer and then validated on an independent validation cohort of 484 H&E-stained WSIs[57]. The model was associated with high levels of concordance, with an area under the receiver operating characteristic (AUROC) of 0.931 and 0.779 for the training and independent cohorts, respectively. A large international consortium trained and validated a CNN model to predict MSI/dMMR from nine cohorts that included 8,343 patients with colorectal cancer across different countries and ethnicities[56]. The CNN model achieved "clinical grade" performance, with an AUROC of up to 0.96, indicating that this AI system can rule out 25-50% of patients for MSI/dMMR testing[56].

CNN models have also been used to predict *EGFR*, *KRAS*, and *STK11* mutations from pathology images with high accuracy[50,51,58–60]. For instance, CNN-based analyses of two large H&E WSI datasets with matched genetic profiling across diverse tumor types were used to predict genetic alterations: The Cancer Genome Atlas (TCGA) dataset was used for model training and the Clinical Proteomic Tumor Analysis Consortium dataset was used for validation[60]. Multiple clinically relevant mutations were predicted (i.e., *PTEN* and *TP53* in endometrial cancer, *KRAS* and *BRAF* in colorectal cancer, and *EGFR* in non-small cell lung cancer [NSCLC]) in both the training and validation sets, demonstrating the potential role of prioritizing patients for confirmatory genetic testing[60].

Another CNN model was developed to predict the molecular classification using H&E WSIs from 2028 patients with endometrial cancer. The patient data were derived from three randomized trials and four clinical cohorts and divided into training and independent validation sets[61]. Using genomic and IHC assessments, patients were classified into one of four prognostic groups: POLE[mut], dMMR, p53 abnormal (p53abn), and no specific molecular profile (NSMP). In the independent validation set, the model achieved class-wise AUROCs of 0.849 for POLE[mut], 0.844 for dMMR, 0.883 for NSMP, and 0.928 for p53abn[61]. Subsequent analysis using ML techniques demonstrated that morphological features including inflammatory, stromal, and tumor cell counts as well as tumor nuclear size and shape were associated with the molecular phenotypes, suggesting the potential for integration into an improved risk stratification system.

Other investigators compared the typical workflow for the diagnosis of prostate cancer (using H&E-stained needle biopsies) with the workflow after introduction of a tool that identifies the need for IHC analysis[39]. They used an ensemble of CNNs to segment tissue from debris and from foci of interest in the H&E-stained WSI and an ML classifier to classify cases as clearly malignant, clearly benign, or ambiguous. This classifier at the time of H&E staining triggered an automated request for IHC in ambiguous cases without waiting for a pathologist's manual review. The AI assistance tool attained 99% accuracy and a 0.99 area under the curve (AUC) on the test data; on a validation set, the average agreement with pathologists was 0.81, with a mean AUC of 0.80. This AI tool to automate IHC requests would, therefore, result in a significantly leaner workflow[39].

These studies indicate that DL computer vision capabilities for predicting molecular characteristics, e.g., genetic mutations and MSI, from H&E-stained WSIs may streamline pathology workflows for known biomarkers.

Table 1 | Overview of Clinical Trials Utilizing AI/ML Methodologies for Cancer Diagnosis, Prognosis, and Treatment Outcomes

First author and reference	Title	Tumor type	Methodology	Outcomes	Lessons learned
Hotta T[13]	Deep learning-based diagnosis from endobronchial ultrasonography images of pulmonary lesions	Lung nodules	CNN, radiomics	A fully automated analysis using EBUS-TBNA cytological WSIs from enlarged mediastinal lymph node demonstrated high precision and sensitivity for lung cancer staging, while performing faster than state-of-the-art baseline models	EBUS-computer-aided diagnosis system may aid diagnosis of malignant lung lesions
Zhang Y[14]	Multimodal imaging under artificial intelligence algorithm for the diagnosis of liver cancer and its relationship with expressions of EZH2 and p57	Liver	CNN	CNN segmentation algorithm predicting liver cancer had high accuracy. A multimodal approach had higher sensitivity, a specific degree of consistency, accuracy, and consistence compared to each modality separately	MUS may have a high clinical application value and accurately predict small liver cancer
Ren C[15]	Clinico-biological-radiomics (CBR) based machine learning for improving the diagnostic accuracy of FDG-PET false-positive lymph nodes in lung cancer	NSCLC	Radiomics, ML	The ML-based DeLong test had the highest predictive accuracy and the lowest false discovery rate among all tested models ($p < 0.05$), both in training and validation sets	The ML-based integration of clinical, biological and radiomics data is associated with increase of the accuracy of lymph node staging in patients with lung cancer, while reducing false positive rates of conventional imaging
Luo X[16]	Automated segmentation of brain metastases with deep learning: A multi-center, randomized crossover, multi-reader evaluation study.	Brain Metastases	DL-based segmentation system using MRI data from 488 patients with 10,338 metastases.	Improved accuracy and efficiency in brain metastasis segmentation.	DL can automate and enhance diagnostic processes in neuro-oncology.
Le Y[17]	CT radiomics analysis discriminates pulmonary lesions in patients with pulmonary MALT lymphoma and non-pulmonary MALT lymphoma.	Pulmonary and Non-Pulmonary MALT Lymphoma	Radiomic analysis using logistic regression, SVM, and KNN with ten selected features.	CNN models had a high accuracy in diagnosing primary pulmonary MALT lymphoma	Radiomics offers significant potential for subtype differentiation in lymphomas.
Arezzo F[18]	A machine learning approach applied to gynecological ultrasound to predict progression-free survival in ovarian cancer patients	Ovarian	ML	ML algorithms were applied to gynecological ultrasound images along with clinicopathological characteristics and created a model to predict 12-month PFS	Prognostic tools may be created by ML-based analysis of ultrasound images combined with other relevant patient data
Zhang K[19]	Using deep learning to predict survival outcome in non-surgical cervical cancer patients based on pathological images	Cervical	DL, digital pathology	Pathomic features from HE-stained images of patients with cervical cancer undergoing chemotherapy and radiation therapy were used for a prognostic model	Pathomic data can be used to create prognostic models
Jiang W[20]	A nomogram based on a collagen feature support vector machine for predicting the treatment response to neoadjuvant chemoradiotherapy in rectal cancer patients	Rectal	SVM	A nomogram incorporating on a support vector machine-based classifier and various clinicopathological characteristics was predictive of response to chemoradiotherapy both in the training and validation set	Collagen features of rectal tumor microenvironment may be associated with response to chemoradiotherapy and predictive models
Sharma A[21]	Development and prognostic validation of a three-level NHG-like deep learning-based model for histological grading of breast cancer	Breast	Digital pathology, CNN	Conventional assessment of NHG and CNN-based grading assessment (predGrade) had similar prognostic performance.	CNN-based model (predGrade) was associated with similar prognostic value with clinical assessment of NHG
Arabyarmohammadi S[22]	Machine learning to predict risk of relapse using cytologic image markers in patients with acute myeloid leukemia post-hematopoietic cell transplantation	AML	ML	Aspirate images from patients with AML and MDS post-allogenic hematopoietic SCT were analyzed using ML-based algorithms to create a prognostic model.	ML-based analysis of aspirate images can predict relapse and prognosis of patients with AML and MDS
Sundar R[23]	Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomized phase III SAMIT trial	Gastric	ML	A random forest model was used to generate a gene signature predicting benefits from adjuvant treatment with paclitaxel in patients with gastric cancer	ML algorithms can be used to identify gene signatures predictive of benefits to various treatment
Christopoulos P[24]	Plasma proteome-based test for first-line treatment selection in metastatic non-small cell lung cancer.	NSCLC	ML algorithm using plasma proteomic profiles for treatment selection.	Improved first-line treatment decisions through biomarker-driven approach.	Plasma proteomics and machine learning enhance personalized treatment in metastatic NSCLC.
Malhaire C[25]	Exploring the added value of pretherapeutic MR descriptors in predicting breast cancer	Breast Cancer		Identified patients at risk of poor response to neoadjuvant chemotherapy.	Multimodal approaches improve predictive accuracy for chemotherapy response.

Table 1 (continued) | Overview of Clinical Trials Utilizing AI/ML Methodologies for Cancer Diagnosis, Prognosis, and Treatment Outcomes

First author and reference	Title	Tumor type	Methodology	Outcomes	Lessons learned
	pathologic complete response to neoadjuvant chemotherapy.		Combining MRI features with clinicobiological predictors including tumor-infiltrating lymphocytes.		
Lv L[26]	Radiomic analysis for predicting prognosis of colorectal cancer from preoperative 18F-FDG PET/CT.	CRC	Developed survival models with clinico-biological and radiomic features.	Radiomics signature integrating PET/CT features and clinical factors achieved a concordance survival index of 0.780 for all CRC stages and 0.820 for stage III CRC, effectively stratifying patients into low-risk and high-risk groups ($P < 0.0001$).	Demonstrated strong correlation between radiomic features and tumor metabolic parameters, indicating that radiomic features can provide valuable insights into CRC prognosis.
Kim C.G[27]	A Phase II open-label randomized clinical trial of preoperative durvalumab or durvalumab plus tremelimumab in resectable head and neck squamous cell carcinoma	HNSCC	Spatial distribution analysis of tumor-infiltrating lymphocytes and high-dimensional profiling of circulating immune cells tracked dynamic intratumoral and systemic immune responses	Preoperative durvalumab with tremelimumab remodeled tumor microenvironment toward immune-inflamed phenotypes, in contrast with durvalumab monotherapy or cytotoxic chemotherapy. High-dimensional profiling of circulating immune cells demonstrated that combination treatment led to a significant expansion and activation of T-cell subsets compared with durvalumab monotherapy.	AI-assisted WSI analysis may reveal changes in tumor microenvironment and circulating immune cells in patients receiving immunotherapy.
Sobottka B[28]	Establishing standardized immune phenotyping of metastatic melanoma by digital pathology	Melanoma	DL	DL algorithms were developed to accurately measure CD8 + T-cell spatial distribution and categorize tumors into clinically relevant immune diagnostic subgroups, including "inflamed", "excluded", and "desert".	The development of a computational diagnostic algorithm that accurately assesses CD8 + T cell densities in various tumor compartments may lead to tumor classification into relevant immune subgroups with clinical implications
Manz CR[29]	Effect of integrating machine learning mortality estimates with behavioral nudges to clinicians on serious illness conversations among patients with cancer: a stepped-wedge cluster randomized clinical trial	All tumor types	ML	Interventions comprising of ML mortality predictions with behavioral nudges that were delivered to oncologists significantly increased the rate of serious illness conversations among patients with cancer and their physicians	ML-based tools can facilitate patient-doctor communication and improve patient care
Guével E[30]	Development of a natural language processing model for deriving breast cancer quality indicators: A cross-sectional, multicenter study	Breast	NLP	Healthcare Quality and Safety Indicators were processed using NLP models to automatically extract indicator elementary variables. The extraction algorithms demonstrate an average accuracy of 76.5%, precision of 77.7% and sensitivity of 71.6%	Although data extraction from unstructured reports can be enabled using NLP, it can be limited by data availability and completeness and algorithm performance.
Ma L[31]	Correlation between AI-based CT organ features and normal lung dose in adjuvant radiotherapy following breast-conserving surgery: a multicenter prospective study	Breast	DL-guided radiation	The proposed DL–based organ feature could accurately predict normal lung dose in patients with breast cancer receiving adjuvant radiotherapy after breast-conserving surgery, possibly reducing the risk for radiation pneumonitis	AI tools can be used to optimize radiation therapy dose.
Chrystall D[32]	Deep learning enables MV-based real-time image guided radiation therapy for prostate cancer patients	Prostate	DL-guided radiation, CNN	A CNN-based real-time image-guided radiation therapy marker was developed and validated in an external cohort, resulting to sensitivity of 98.31% and specificity of 99.87%.	CNN can successfully identify implanted prostate markers

AI Artificial Intelligence, *AML* Acute Myeloid Leukemia, *CBR* Clinico-biological-radiomics, *CNN* Convolutional Neural Network, *CRC* colorectal cancer, *DL* Deep Learning, *EBUS* Endobronchial Ultrasound, *HE* Haematoxylin & Eosin, *HNSCC* Head And Neck Squamous Cell Carcinoma, *MDS* Myelodysplastic Syndromes, *ML* Machine Learning, *MUS* Multimodal Ultrasound, *NHG* Nottingham Histological Grade, *NLP* Natural language processing, *NSCLC* Non-Small Cell Lung Cancer, *SCT* stem cell transplant, *SVM* Support Vector Machine, *WSI* Whole Slide Imaging

AI-based biomarker prediction from H&E-stained WSIs is limited by the following: only molecular biomarkers that have an impact on tissue morphology can be identified; the sensitivity and specificity of AI-based mutation identification is suboptimal; and concordance with validated methodologies is limited (owing to limited tumor tissue availability, poor DNA quality, inaccuracy of laboratory procedures, and lack of personnel experience or other resources). In order for AI-based biomarker prediction from H&E-stained WSIs to be applied in clinical practice, extensive validation in external datasets and within clinical trials is needed.

The use of AI to predict novel prognostic and predictive biomarkers from H&E-stained WSIs. Challenges associated with the complexity and heterogeneity of the immune tumor microenvironment and predictive/prognostic biomarkers may be overcome by computational pathology technologies[62–65]. The Multiomics Multicohort Assessment platform analyzed H&E-stained WSIs from patients with early-stage colorectal cancer using large publicly available datasets, such as TCGA, that included digital H&E-stained WSIs annotated with sequencing and clinical data[62]. The investigators employed CNNs and vision transformers to investigate whether DL analysis of H&E-stained WSIs could predict clinical and molecular profiles of interest. The model accurately predicted clinical outcomes, including overall and disease-free survival, as well as molecular aberrations including copy number alterations, expression levels of key genes in cancer development, MSI, *BRAF* mutation, and CpG island methylator and consensus molecular subtypes[62].

A prognostic model for prostate cancer was developed that incorporated CNN analysis of prostate biopsy H&E-stained WSIs with six clinical variables (combined Gleason score, Gleason primary, Gleason secondary, T-stage, baseline PSA, age) from 5654 patients from the Radiation Therapy Oncology Group prostate cancer studies[63]. This model was shown to have better prognostic accuracy than the commonly used National Comprehensive Cancer Network (NCCN) risk-stratification tool[63]. Similar multimodal DL approaches have been used to predict outcomes for patients with gliomas[64] and high-grade serous ovarian cancer[65]. A local-global graph-based distillation (ALL-IN) model combining both local and global histological features using a graph-based neural network improved stratification of patient risk groups, with clinical utility[66].

Investigators developed a CNN tumor-infiltrating lymphocyte (TIL) "analyzer" to identify three immune phenotypes (IPs)—inflamed, immune-excluded, and immune-desert—based on the concentrations of TILs in tumor epithelium and tumor stroma on H&E-stained WSIs[67]. The inflamed IP (high TIL concentration in tumor epithelium) was associated with higher response rates and longer PFS in studies of immune checkpoint inhibitor therapy in patients with NSCLC. The TIL analyzer provided prognostic insight in addition to the PD-L1 TPS in the subset of patients with a TPS of 1%–49%. The 42.5% of patients with an inflamed IP had a 22% response rate compared with a response rate of only 3.9% in patients with immune-excluded or immune-desert IPs[67].

These studies demonstrate that DL approaches based on H&E image analysis alone or combined with clinical data hold promise for improving prognostic and predictive biomarkers in precision oncology.

Challenges in the implementation of AI/ML tools in digital pathology

The performance of AI/ML tools can complement that of medical doctors in the interpretation, analysis, and conclusions derived from large-scale datasets. The integration and analysis of large-scale datasets such as genomic, radiomic/radiogenomic, digital pathology, real-world, and EHR datasets requires advanced computational tools and increased power, owing to their complexity and heterogeneity. The TCGA includes more than 10,000 digital pathology images from patients with diverse tumor types, along with associated clinicopathological and genomic data (https://www.cancer.gov/tcga). The Virchow2G Pathology Dataset includes over 3 million pathology slides from 225,000 patients across 45 countries and was used to train Virchow2G, a large pathology model[68].

The Cancer Imaging Archive comprises de-identified medical images of cancer that are associated with patient outcomes, treatment, and genomic data[69]. These large-scale datasets present challenges related to the management and storage of large volumes of data, increased variety of data sources and formats, assessment of batch effects, high processing power requirements, and tool integration, along with relevant feature selection, which is often hindered by nonlinear associations of different features and inter-tumor and intra-tumor heterogeneity. AI/ML algorithms enable the extraction of clinically relevant features from these datasets, providing useful insights that could not be identified by traditional methods or human intelligence.

Digital pathology, while transformative for the application of precision oncology, poses several challenges. The generation of "big data" requires efficient data management and storage systems, and interoperability issues associated with the lack of compatibility of different digital pathology systems across platforms and institutions limit data sharing and integration. The regulatory and legal framework for the use of digital pathology is evolving, and concerns regarding data privacy and the need for standardization of practices should be addressed. In addition, quality control and methodology validation, along with pathologist training, are critical for the application of digital pathology in clinical practice. The transition from traditional to digital workflows may be challenging, requiring time and adaptation. Finally, the increased costs associated with the integration of digital pathology, including scanning equipment, specialized software, data storage, technology infrastructure, and extensive physician training, may be a significant barrier for smaller institutions.

Multiplex, single-cell, and digital spatial analyses. AI/ML tools have the potential to analyze the emerging complex and highly dimensional measurements of disease, offering deeper understanding of tumor biology, including the interaction of the tumor microenvironment with the tumor. They can help analyze results derived from digital pathology *multiplex* platforms that measure multiple analytes in a single sample, such as gene expression at the protein (IHC, immunofluorescence, or imaging mass cytometry) and mRNA (bulk or single-cell RNA sequencing) levels. AI/ML tools are increasingly employed for the characterization of individual cells using protein, DNA, RNA, and metabolite analysis to pinpoint single-nucleotide mutations[70–73] and for the investigation of epigenomic phenomena such as DNA methylome[74–76], ChIP-seq analysis, and chromatin accessibility data[77,78].

Applying AI algorithms, various tissue types can be classified based on their spatial characteristics (texture, shape, and color). The spatial distribution of cancer and neighboring cells can be combined with other clinicopathological data to establish prognostic and predictive algorithms. For instance, imaging mass cytometry was applied to evaluate the tumor and immunological landscape of tissue samples from 416 patients with NSCLC and to assess a prognostic model[79]. Investigators demonstrated that CNN-based spatial analysis of immune lineages and activation status identified five markers (CD14, CD16, CD94, αSMA, and CD117) that correlated with OS[79].

In another study, imaging mass cytometry-labeled brain tumor biopsies were used to create high-dimensional maps of the brain tumor microenvironment[80]. CNN algorithms enabled fully automated high-throughput segmentation and identification of individual cells across diverse tissues. Differences in the tumor immune landscapes between patients with high-grade glioma and brain metastasis were observed. Spatial cellular neighborhoods (CNs) that were associated with OS were identified in patients with glioblastoma. Furthermore, CNs enriched in M1-like monocyte-derived macrophages were associated with improved OS, highlighting the value of spatial cellular relationships and showing the complexity of tumor CNs[80]. Others used multiplexed ion beam imaging by time-of-flight (MIBI-TOF) with a CNN segmentation tool to evaluate in situ expression of 36 immune-related proteins in patients with triple-negative breast cancer and to define the tumor-immune microenvironment, including identification of CNs[81].

Other researchers developed a weakly supervised (e.g., not requiring manual expert annotation) DL framework to identify tumor-immune interrelations and CNs and to predict which patients with low-risk early-stage endometrial cancer have a higher risk of recurrence[82]. Using multiplexed immunofluorescence of tissue microarrays from tumor samples for the simultaneous visualization and quantification of CD68+ macrophages, CD8 + T cells, FOXP3+ regulatory T cells, PD-L1/PD-1 protein expression, and tumor cells, they trained and validated a multilevel interpretable DL framework (using a CNN for patch feature extraction, a graph neural network to capture CNs and tissue areas, and a multilayer perceptron for recurrence risk classification) to predict the risk of recurrence. This model achieved an AUROC of 0.90, and predictions resulted in concordance for 96.8% of cases. The authors concluded that the model could assess the risk of recurrence in this study population, outperforming current prognostic factors, including molecular subtyping[82].

Another promising approach combines AI-driven image analysis of cellular phenotypes with automated single-cell or single-nucleus laser microdissection and ultra-high-sensitivity mass spectrometry. This approach links protein abundance to cellular and subcellular phenotypes while preserving spatial context, offering the potential to elucidate pathways that change in a spatial manner as cancer progresses[83].

In addition, RNA sequencing (RNAseq) plays a crucial role in multiplex analyses, providing a comprehensive view of gene expression profiles within the tumor microenvironment. The integration of RNAseq data with AI/ML tools allows for the identification of novel biomarkers and gene signatures that are pivotal for understanding tumor biology and patient outcomes. Investigators have used an autoencoder, an unsupervised DL methodology that utilizes input data to create representative features, to regenerate output data and integrate DNA methylation, RNAseq, and miRNAseq data from patients with colorectal cancer[84]. This approach enabled the identification of a subgroup of patients with improved OS. Another study highlighted that the clustering algorithms applied to RNAseq data can uncover distinct gene expression patterns that correlate with specific tumor characteristics, thereby facilitating the identification of potential therapeutic targets[85]. This synergy between RNAseq and AI-driven analyses not only enhances the characterization of tumor-immune interactions but also supports the development of prognostic models that can predict patient responses to therapies. By leveraging the high dimensionality of RNAseq data in conjunction with spatial and multiplex imaging techniques, researchers can gain deeper insights into the complex interplay between tumor cells and their microenvironment, ultimately advancing precision medicine approaches in oncology.

In summary, advanced multiplex imaging technologies coupled with AI analytics enable a deepened understanding of tumor-immune interactions in the tumor microenvironment and may enable the discovery of novel biomarkers and therapeutic targets.

Digital radiology (radiomics)

In the past decade, the field of medical image analysis has grown exponentially, with an increased number of pattern recognition tools and larger data sets. Radiomics refers to the high-throughput mining of quantitative image features from standard-of-care medical imaging that enables data to be extracted and applied within clinical decision support systems to identify complex patterns and trends for improving diagnostic, prognostic, and predictive accuracy[86]. This approach expands the utility of radiologic data beyond medical images that are simply visual aids for human interpretation[82].

Digital medical images are converted into mineable high-dimensional quantitative data in a matrix format where each element, known as a voxel, corresponds to a small section of the body. These voxels contain x-ray attenuation values directly proportional to the density of the material being scanned, with a total range of more than 4096 intensities, while only a small fraction of these intensities can be perceived by humans. The limited discriminatory capacity of the human eye suggests the potential for DL methods[87]. Quantitative radiomic features, measured or mathematically

transformed, representing intensity, geometry, and texture may reflect aspects of the tumor phenotype and microenvironment that can predict clinical outcomes and support clinical decisions.

Image segmentation involves partitioning an image into meaningful regions, which is essential for accurately identifying tumors and organs at risk in radiation oncology. Accurate segmentation is crucial for treatment planning, as it directly impacts the precision of radiation delivery. Traditional manual segmentation is not only time-consuming but also prone to inter-observer variability, which can lead to inconsistent results.

For instance, the BRATS (Brain Tumor Segmentation) challenge, an annual international competition focused on brain tumor segmentation, has been instrumental in driving advancements in this field. This challenge encourages the development of innovative segmentation algorithms and fosters collaboration among researchers, leading to improved methodologies and performance benchmarks. A research group introduces a weakly supervised approach to pan-cancer segmentation, showcasing the potential of AI/ML to tackle complex segmentation tasks, even with limited annotation[88]. Their method leverages slide-level annotations to train segmentation models, demonstrating that effective tumor segmentation can be achieved without extensive pixel-level labeling, which is often a bottleneck in clinical practice[88]. Many investigators have reported on segmentation algorithms for various organs, such as the liver[89], brain[90], pancreas[91], and prostate[92,93]. Guidelines for the development, clinical validation, and reporting of AI models in radiation therapy have been developed by the European Society for Therapeutic Radiation Oncology and the American Association of Physics in Medicine for the standardization of this approach[94].

Investigators used statistics and ML (Least Absolute Shrinkage and Selection Operator) to develop a radiomic model to predict TIL density, as determined from AI-powered analysis of H&E-stained WSIs, using the same technology as previously described[67], and baseline CT imaging from a training cohort of 220 patients with NSCLC treated with immunotherapy[95]. The final ML-based TIL-prediction model included only two features, both indicative of intralesional texture heterogeneity, and demonstrated that high predicted TIL density (≥ median) was associated with longer PFS compared to low predicted TIL density (median, 4.0 months vs. 2.1 months, $p = 0.002$) when applied to a 294-patient validation cohort. TIL density was significantly associated with PFS independent of PD-L1 status, and patients with high TIL density and high PD-L1 (TPS ≥ 50%) had the longest PFS compared with patients with low TIL density and/or PD-L1 TPS[95].

In addition, radiomics has been used to predict immunotherapy outcomes[96]. For instance, investigators have evaluated CT imaging data from 54 patients with hepatocellular carcinoma treated with immunotherapy using nine ML and two ensemble learning techniques to construct predictive models[97]. The models were validated in an external set comprising 29 patients; selected ML models were shown to accurately predict the short-term efficacy of immunotherapy in patients with hepatocellular carcinoma[97]. Other investigators used radiological images annotated with clinical and outcome data from 2552 patients to develop ML models to predict OS in patients with head and neck cancer and validate the models in three external cohorts comprising 873 patients[98]. Among 12 different models, one achieved the highest prognostic accuracy using multitask learning on clinical data and tumor volume. However, the results demonstrated significant decreases in model performance, and could not be validated in the external datasets[98].

Other investigators constructed and validated a sub-regional radiomics model based on a support vector machine algorithm using 1896 features from each tumor sub-region, (5688 features per sample) from 264 patients with NSCLC[99]. In the validation set, the model demonstrated improved accuracy in predicting immunotherapy response compared to conventional radiomics, tumor mutational burden (TMB), or PD-L1[99].

In another study, an ML (random forest) prognostic radiomic model was developed using CT images from patients with advanced melanoma who participated in pembrolizumab multicenter clinical trials[100]. The model achieved a high AUC for OS estimation in the validation set, suggesting that

this tool could be used for clinical decisions[100]. Based on radiomics features, other investigators used pre-operative CT images from 127 patients with NSCLC from TCGA to construct a TMB prediction model[101]. Three radiomics features (flatness [shape of original feature], autocorrelation [GLCM], and minimum [first order of wavelet features]) were found to be associated with TMB levels and were significantly different between the high- and low-TMB groups[101].

Additional ML radiomics models have been developed to identify patients who may benefit from immunotherapy (e.g., patients with melanoma, NSCLC, or breast cancer)[100–102]. The above examples employed ML techniques to assist in the analysis of non-ML-derived, classical, "hand-crafted" (i.e., human-defined) radiomic features. Recent investigations employing CNNs for DL of features may outperform approaches using handcrafted radiomic features[103,104]. For example, transformers and novel architecture methodologies have shown promising results in improving feature extraction and diagnostic accuracy in medical imaging tasks[105,106].

An ML model trained on dual-energy CT radiomics (DECT) was shown to be superior to standard CT imaging and enabled quantification of iodine and fat concentrations in lesions, in addition to visual inspection[107]. The application of DECT to an ML-based radiomics model significantly improved immunotherapy response prediction for patients with stage IV melanoma compared to standard CT imaging[107].

The application of AI/ML algorithms has been shown to improve or surpass the performance of physicians in cancer diagnosis and staging[108–110]. In one study, an AI model trained on 506 CT images exhibited better diagnostic accuracy in distinguishing benign vs. malignant pulmonary nodules compared to different groups of physicians[108]. In another study, an AI-based model developed and validated on 170,230 mammography images demonstrated higher diagnostic performance in terms of breast cancer detection compared to radiologists[109]. However, with the addition of AI, the performance of radiologists significantly improved.

In summary, use of AI/ML techniques in radiomics analysis can transform medical imaging data into quantifiable variables that may be used as noninvasive prognostic and predictive biomarkers for response to treatment, overcoming the limitations of tissue-based analysis for clinical decision-making. However, these preliminary data warrant validation in larger patient cohorts.

Challenges regarding the use of AI/ML techniques in radiomics include the following: lack of prospective analyses of imaging data; lack of evaluation of radiomics within prospective clinical trials or standardized and homogeneous frameworks; a limited number of studies with independent validation of the results and their interpretability; and a lack of training and knowledge of physicians on radiomics. Data reproducibility across different datasets is hindered by various methodological approaches, including variability in imaging protocols among different hospitals, heterogeneity in patient populations, preprocessing (image normalization, noise reduction, and image segmentation), feature selection, and model training[111]. Developing multicenter studies assessing the standardization of protocols and workflows in medical imaging is important to ensure reproducibility and applicability across institutions.

Molecular medicine

The exponential growth of techniques to assess "omics" data, including next-generation sequencing (NGS) techniques, has contributed to the identification of novel prognostic and predictive biomarkers and drug targets. A challenge in genomic analysis using NGS is the annotation of molecular alterations and variant calling, e.g., identifying the differences between the analyte sequence (patient's sample) and the reference sequence[112]. This process is prone to errors, ranging from 0.1% to 10%, and has important clinical consequences. Variant callers based on ML models (logistic regression; hidden Markov models; naïve Bayes classifiers), such as the Genome Analysis Toolkit, had less than optimal accuracy, even on short-read sequencing technologies such as Illumina with 75-250 bases, and were poorly generalized to the newer long-read NGS technologies, such as Pacific Biosciences with 15,000 bases and Oxford Nanopore with up to 1

million bases[113,114]. A major step forward was the implementation of CNNs in variant calling as exemplified by DeepVariant[113]. This model outperformed all other existing tools, winning the highest performance in an FDA-administered variant calling challenge. Furthermore, this model performs well on both short-read and long-read whole genome and exome sequencing technologies and generalizes even to other mammalian species[113].

AI/ML tools have also been used to analyze large-scale epigenomic datasets to identify patterns associated with specific tumor types[115], which can serve as biomarkers for early detection[75], accurate diagnosis[116], and prediction of patient outcomes[117]. By analyzing large-scale genomic and epigenomic data sets, AI can help discover novel epigenetic drugs, repurpose existing drugs, identify potential candidates that target specific epigenetic modifications[118], and develop predictive models with integration of epigenomic, clinical, and patient outcomes data[70,71].

AI/ML tools have also been used for the analysis of the output of proteomic measurement techniques. A "sample-to-data" roadmap for integrating AI/ML throughout the proteomic workflow has been suggested[73]. In another study, an AI algorithm was developed to identify protein interaction networks for individual patients based on their proteomic profiling data[119], indicating that interaction networks may be accurately reconstructed, representing an advancement over standard methods[119].

Integrative (multimodal) analyses

Most applications of AI/ML in precision oncology represent "narrow" tasks using one data modality such as pathology, radiology, or molecular sequencing data. However, oncologists integrate all relevant available modes of data when evaluating patients. The task of modality conversion is central to advancing AI in network medicine applications[120–123]. Modality conversion involves transforming data from one form to another, which is crucial for enabling AI to mimic human-like sensory integration and interpretation. One example in the field of radiation therapy is the use of DL tools for the generation of synthetic CT images from magnetic resonance images to aid in radiation therapy planning[124]. Transformer-based text, vision, and speech models can facilitate these conversions. Multiomic or panomic technologies using AI/ML/DL tools may improve the discovery of molecular biomarkers[125]. Emerging AI methodologies can drive the progress in network medicine, ultimately improving patient outcomes and uncovering novel therapeutic targets[120,121].

The development of multimodal AI models incorporating all relevant sources of data—eventually including biosensor (devices that continuously detect and measure physiologic or environmental parameters to assess specific biomarkers), social determinants, and environmental data—is becoming potentially feasible[126]. Investigators developed a multimodal classifier to predict response to PD-L1 blockade in patients with NSCLC[127] that included the clinical, pathological, radiomic, and genomic characteristics of 247 patients treated at a single center. Radiomic features were extracted using classical radiomics techniques; PD-L1 tumor cell expression was assessed as the standard TPS; a CNN model was also used to develop an automated PD-L1 classifier on digital PD-L1-stained WSIs; and genomic analysis assessed somatic mutations, copy number alterations, and fusions in 341-468 genes most associated with cancer and TMB. Clinical data included neutrophil-to-lymphocyte ratio, pack-years smoking history, age, albumin, tumor burden, presence of brain and liver metastases, tumor histology, and scanner parameters. An attention-based DL model was developed that could account for non-linear relationships across the input modalities. The model was able to predict objective responses better than any modality separately or linearly combined and led to enhanced separation of Kaplan-Meier survival curves (indicating potential as a useful biomarker for longer-term outcome). Analysis of the model revealed that all data modalities (radiomics, genomics, and pathology) contributed to the prognostic classification success[127]. Frameworks, such as Prototypical Information Bottlenecking and Disentangling, were used to address redundancy issues in multimodal data, thereby improving cancer survival

predictions[128]. In summary, the application of AI/ML algorithms to integrated medical multimodal data has great promise and will depend on the assembly of large, well-annotated, multi-institutional training datasets[127].

Large language models and generative AI
Many useful applications in the field may result from AI advances in NLP, especially with the development of LLMs. The advent of LLMs with a user interface, which enables communication between the AI system and a human using natural language, has facilitated the emergence of "generative" AI, i.e., technology that can generate text, images, or other data (video, sound) based on features learned from input training data[129].

After training on big data, LLMs can perform various tasks, including summarization, translation, text completion, and imaginative writing[130]. LLMs have been leveraged to facilitate decision support for patients with cancer[131–135]. A panel of clinicians evaluated the responses of Almanac, an LLM augmented with retrieval capabilities from curated medical sources, to clinical questions including medical guidelines and treatment recommendations[134]. Almanac's responses to 314 clinical questions were better than the other LLMs (ChatGPT-4, Bing, and Gemini) that were not augmented with medical data[134].

The use of Med-PaLM Multimodal, a multimodal generative LLM finetuned on medical data, was associated with high performance across diverse tasks including responses to medical questions, interpretation of mammography and dermatology images, radiology report generation and summarization, and genomic variant calling. The application of this multimodal LLM indicates the potential for the broader use of medical AI systems[136].

Other applications of LLMs include mining of EHRs to identify clinically relevant data, such as treatment-related adverse events, and to support insurance reimbursement[137]. The LLM GatorTron was successful in recognizing adverse events attributed to certain drugs[137]. If validated, this approach may improve patient care[138].

However, the application of LLMs should be interpreted with caution because it is associated with challenges. One example is the poor performance of an LLM chatbot (ChatGPT) in terms of providing treatment recommendations concordant with NCCN guidelines[139]. High rates of discordant responses and "hallucinations" (e.g., responses not related to any recommended treatment) were identified in 13 (12.5%) of 104 ChatGPT outputs. These "hallucinations" have been previously described as a critical issue with AI chatbots[140]. Other challenges related to the use of LLMs are accountability, research integrity, and data security. In summary, LLMs cannot be incorporated into clinical practice at this time. Thorough clinical validation using stringent criteria is required from developers to ensure high rates of accuracy in terms of generative AI predictions and responses, and clinicians should be aware of their limitations.

FDA-approved AI/ML-enabled medical devices
As of December 20, 2024, the FDA has approved 1016 AI/ML-enabled medical devices that are authorized for marketing in the United States[141]. Specific examples where AI has successfully impacted clinical outcomes, underscoring the real-world applicability, are listed in Table 2.

Ethical and regulatory aspects of AI deployment in precision oncology
The rapid evolution of AI in precision oncology necessitates thorough ethical and regulatory considerations related to biases associated with data, model transparency, and accountability.

Data bias. One of the major concerns is data bias, as AI models were often trained with non-representative or biased datasets[142]. Unintentional existing biases within the healthcare system may contribute to treatment inequities among marginalized racial and ethnic groups if the training data do not adequately represent these populations[143]. Biases in healthcare research and public health databases may mislead AI outputs, which

may negatively affect treatment recommendations and patient outcomes[144,145].

Model transparency and trust. The complexity of AI algorithms is often associated with lack of transparency, which may result in healthcare professionals feeling uncertain about the reliability of AI applications[146]. Clinicians may hesitate to rely on AI recommendations due to the "black box" nature of many models[147,148]. Explainable AI (XAI) methods are essential for building trust in AI recommendations, helping users to understand the reasoning behind the suggestions, providing transparency, and boosting confidence in the decisions made[149,150].

Accuracy and reliability. The development of clinical decision support systems is ongoing, and these systems cannot yet be utilized because of the inaccuracy and unreliability of AI predictions[147,148]. Rigorous clinical validation, standardization, and real-world testing are essential before deployment. Transparency about model limitations and monitoring of performance post-deployment are critical to maintaining clinical safety.

Accountability. As AI systems become integrated into healthcare, issues regarding accountability and liability that may adversely affect a patient's health should be addressed. A clear guideline that delineates the responsibilities of AI developers, healthcare providers, and institutions is necessary[151]. Effective post-market surveillance mechanisms to monitor the performance of AI systems after deployment and ensure that they continue to operate within ethical and clinical standards should be implemented by the regulatory agencies[152].

Data privacy and ethical use. AI systems require extensive patient data, raising privacy and ethical concerns regarding consent, ownership, and secondary use[147,153]. Transparent policies governing how patient data are collected, stored, and shared that align with regulations such as the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) are essential to ensure ethical standards and respect for patient rights[147,153]. As the field evolves, continuous collaboration among stakeholders, including ethicists, legislators, and medical practitioners, is necessary to advance the ethical and effective integration of AI in healthcare. Harmonization of policy and practice are essential components of the implementation of AI/ML in the clinical workflow.

Data privacy and inter-institutional collaboration in AI-driven oncology
The advancement of AI applications in oncology requires extensive, diverse datasets for model training and validation. However, sharing sensitive patient data across institutions presents significant privacy, regulatory, and ethical challenges. The data that was once a byproduct of clinical research is increasingly becoming a resource[154]. Data management includes ensuring the safety, accessibility, and accuracy of the data. Guidelines and processes to access and curate data and alignment with regulatory and compliance departments are essential elements of data management[155].

Federated learning (FL) has emerged as a transformative solution that enables multi-institutional collaboration without compromising patient privacy. This approach allows AI models to be trained across multiple institutions while keeping patient data securely within their original locations[156]. In FL, instead of centralizing data, the training algorithm travels to each institution's secure environment, learns from local data, and only shares model parameters rather than raw patient information[156,157].

A critical component of successful multi-institutional collaboration is data harmonization. Modern platforms implement standardized clinical data harmonization pipelines that enable FL including Fast Healthcare Interoperability Resources standards and automated data transformation workflows[158]. Furthermore, FL architectures are designed to comply with major privacy regulations, including GDPR and HIPAA, ensuring that data

Table 2 | Selected examples demonstrating AI successfully impacting clinical outcomes, underscoring real-world applicability

Author/Organization	Year	Tumor Type	Methods Used	Results	How AI Changes the Practice
Liu et al.[191]	2017	Breast Cancer	CNN-based tumor detection and localization in gigapixel pathology images using Inception (V3) architecture with multi-scale inputs.	Detected 92.4% of tumors at 8 FP/image, exceeding previous methods (82.7%) and human pathologists (73.2%).	Improved sensitivity and reduced false negatives in breast cancer metastasis detection, aiding pathologists in diagnosis.
Esteva et al.[192]	2017	Skin Cancer	CNN trained on dermoscopic images for lesion classification.	Matched performance of 21 board-certified dermatologists in identifying malignant vs benign lesions.	Enabled wider access to dermatological expertise via AI-powered applications.
Lay N et al.[193]	2017	Prostate Cancer	Random forest classifiers leveraging spatial, intensity, and texture features from MRI sequences (T2W, ADC, B2000).	Achieved an AUC of 0.93, outperforming the SVM-based CAD approach (AUC of 0.86) on the same test data.	Improved prostate cancer detection accuracy by leveraging instance-level weighting and enhanced feature extraction.
Zhang C et al.[194]	2019	Lung Cancer (Pulmonary Nodules)	3D CNN; CT Imaging; Open-source and multicenter datasets	Sensitivity: 84.4%; specificity: 83%; superior to manual assessment; effective for nodules <10 mm and 10–30 mm	AI assisted radiologists by providing objective, accurate, and timely detection/classification of nodules, improving diagnostic efficiency.
McKinney et al.[195]	2020	Breast Cancer	CNN models trained on large datasets of mammograms to detect malignancies, leveraging transfer learning.	Reduced false positives by 5.7% and false negatives by 9.4% compared to radiologists	Improved early detection, reducing unnecessary biopsies, and missed diagnoses.
Exscientia[196]	2021	General (Drug Discovery)	AI-driven molecular design using generative algorithms and reinforcement learning to identify potential cancer therapeutics.	Reduced drug design timeline significantly, with the first AI-designed drug entering clinical trials in record time.	Accelerated the drug development process and personalized therapy options by identifying targeted molecular candidates.
Zhu S et al.[197]	2021	Breast, Lung, Prostate, Thyroid, Bone Cancers	Analysis of FDA-cleared AI/ML devices, primarily using rule-based algorithms for diagnostics and DL for radiation planning.	52 out of 343 devices (15.2%) were oncology specific. Since 2016, 94.2% of these were approved. The majority (96.2%) cleared by 510(k) pathway. Diagnostic devices mainly targeted breast (66.7%), lung (14.3%), prostate (6.3%), thyroid (6.3%), and bone (6.3%) cancers. Therapeutic devices (30.8%) were primarily for radiation planning.	Enhanced diagnostic accuracy and efficiency in detecting oncologic pathologies such as breast cancer. Improved treatment precision with AI-driven radiation planning, automating organ segmentation for radiotherapy.
GI Genius[198]	2021	Colon Cancer	GI Genius: AI device for real-time lesion detection during colonoscopy using deep learning algorithms.	FDA-authorized; improved adenoma detection rates during colonoscopy.	Assisted clinicians in real-time detection of colorectal polyps, enhancing early cancer detection.
Optellum[199]	2021	Lung Cancer	AI-driven software leveraging probabilistic ML for early lung cancer detection from CT scans.	Received FDA clearance for software aiding in the early detection and optimal treatment of lung cancer.	Enhanced early diagnosis, improving patient outcomes through timely intervention.
Paige.AI[200]	2021	Prostate Cancer	The Paige Prostate system employs a deep learning algorithm to analyze whole slide images of prostate needle biopsy slides, providing binary classifications and localizing the highest-probability cancer regions using annotated datasets for training and validation.	The system improved diagnostic sensitivity by 7.3% for cancer cases, increased specificity by 1.1%, and enabled pathologists to identify overlooked cancer areas, enhancing overall diagnostic accuracy.	AI complements pathologists by highlighting suspicious regions, improving detection of subtle cancers, reducing diagnostic variability, and optimizing focus areas for review, leading to more accurate and efficient pathology workflows.
Ye M et al.[201]	2022	Lung Cancer	AI-enhanced classifier combining DL for imaging and liquid biopsy for early diagnosis.	Improved diagnostic accuracy in early-stage lung cancer detection, enhancing sensitivity and specificity compared to traditional methods.	Combined AI and liquid biopsy approach offers a non-invasive, accurate method for early lung cancer diagnosis, potentially improving patient outcomes.
Esteva A et al.[63]	2022	Prostate Cancer	Multi-modal DL model integrating clinical data and digital histopathology from prostate biopsies. Model: trained, validated using data from 5 Phase III trials (median follow-up, 11.4 years)	Demonstrated superior discriminatory performance compared to the NCCN risk stratification tools; relative improvement (range, 9.2% to 14.6%) in predicting long-term, clinically relevant outcomes.	Enhanced prognostication by allowing oncologists to computationally predict patient-specific outcomes, facilitating personalized therapy decisions in prostate cancer treatment. This AI-based tool is scalable and can be implemented globally in clinics equipped with digital scanners and internet access.
ProstatID[202]	2022	Prostate Cancer	ProstatID: AI software using ML on prostate MRI data for cancer detection.	Received FDA clearance for improving accuracy and speed of prostate cancer detection.	Enhanced MRI diagnostics, leading to earlier and more accurate prostate cancer detection.
NIH Algorithm-Based Tool[203]	2022	Prostate Cancer	AI-driven software trained on annotated pathology datasets for cancer detection	FDA cleared; demonstrated improved detection accuracy.	Automated and enhanced prostate cancer diagnosis in clinical settings.
Liu M et al.[204]	2023	Lung Cancer	Meta-analysis of multiple AI techniques, including supervised learning (CNNs, random forests) for CT scan analysis.	AI systems demonstrated high accuracy in diagnosing lung cancer, with pooled sensitivity and specificity rates indicating reliable performance.	Enhanced early detection and diagnostic accuracy in lung cancer, potentially leading to improved patient outcomes.

Table 2 (continued) | Selected examples demonstrating AI successfully impacting clinical outcomes, underscoring real-world applicability

Author/Organization	Year	Tumor Type	Methods Used	Results	How AI Changes the Practice
Spratt DE et al.[205]	2023	Prostate Cancer	An ensemble model using digital pathology images combined with clinical trial data to predict distant metastasis risk.	In the validation cohort (14.9 years follow-up), ADT significantly improved time to distant metastasis for model-positive patients (34%). No benefit for model-negative patients.	AI enabled patient-specific predictions, improving decision-making for targeted use of ADT in prostate cancer treatment.
Thirona[206]	2024	Lung Cancer	AI-based software analyzing lung CT images to detect cancer and other lung diseases.	Received FDA clearance for lung analysis software to assist in diagnosing lung diseases, including cancer.	Improves diagnostic accuracy and efficiency in lung cancer detection.
Avenda Health[207]	2024	Prostate Cancer	Unfold AI: AI tool using deep learning on prostate MRI scans for cancer detection.	Achieved 84% accuracy in detecting prostate cancer, outperforming doctors' 67% accuracy.	Enhances diagnostic precision, leading to more targeted and effective treatments.
Guardant Health[208]	2024	CRC	AI-driven blood-based screening test (Shield) for detecting cancer biomarkers.	FDA-approved; detected 83% of colorectal cancers in clinical studies.	Provides a less invasive and more accessible screening option, potentially increasing screening rates.
ArteraAI[209]	2024	Prostate Cancer	AI prognostic tool leveraging genomic data to predict outcomes in localized prostate cancer.	Included in NCCN Clinical Practice Guidelines as a predictive test.	Assisted in personalized treatment planning by predicting patient outcomes.
DermaSensor[210,211]	2024	Melanoma, BCC and SCC	AI-powered device using light reflectance for non-invasive skin cancer diagnosis.	FDA-approved device for detection of all three major skin cancers.	Expanded access to accurate skin cancer screening in dermatology and primary care settings.

ADT Androgen Deprivation Therapy, AI Artificial Intelligence, AUC Area Under the Curve, BCC Basal Cell Carcinoma, CAD Computer-Aided Diagnosis, CNN Convolutional Neural Network, CRC Colorectal Cancer CT Computed Tomography, DL Deep Learning, FDA Food and Drug Administration ML Machine Learning, MRI Magnetic Resonance Imaging, NCCN National Comprehensive Cancer Network, SCC Squamous Cell Carcinoma, SVM Support Vector Machine.

remain within institutional boundaries and that there is no direct sharing of protected health information[156,157].

This privacy-preserving approach to multi-institutional collaboration represents a paradigm shift in how healthcare data can be utilized for research while maintaining the highest standards of patient privacy and data security. This is especially important for the future of AI in oncology to ensure that training datasets are large, diverse, and inclusive of low-frequency "rare" cancers, thereby ensuring generalizability and clinical utility.
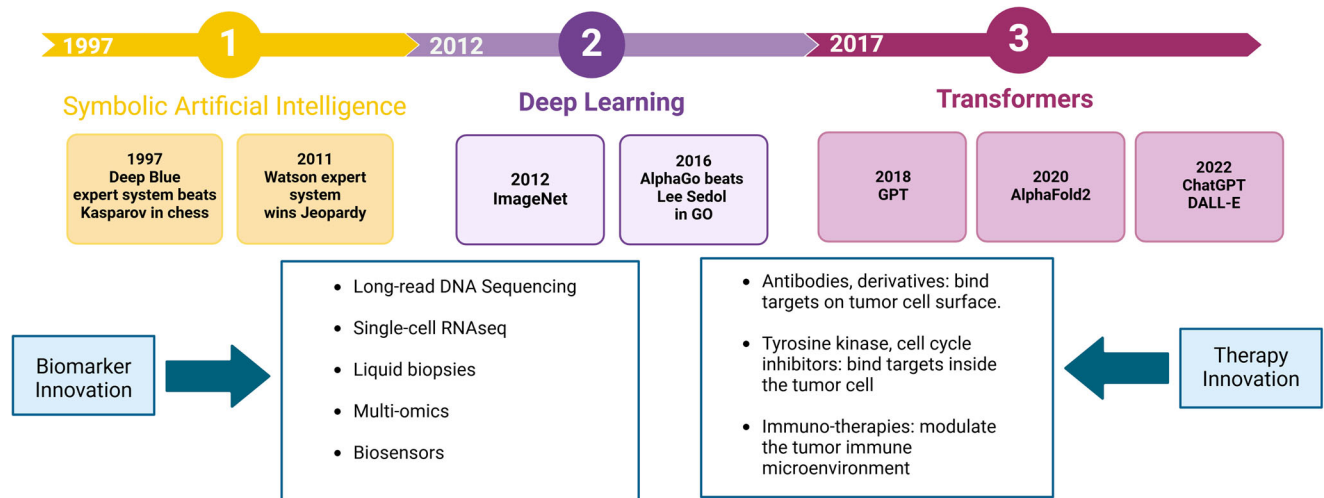
Future directions and emerging trends

Biosensors are devices or platforms that continuously detect and measure physiologic or environmental parameters to assess specific biomarkers associated with diverse diseases, including cancer. They comprise a biological sensing component and a transducer responsible for converting the identified signal into a quantifiable output. The combination of AI/ML with biosensors for the real-time continuous monitoring of physiologic parameters may provide new clinically relevant insights into the early diagnosis, prognosis, and treatment of cancer. AI-based biosensors are being evaluated in diverse tumor types to improve early detection[159–162], diagnosis[163–165] and treatment outcomes[166,167]; and they should be further validated in large studies. In addition, several biosensors continuously measure parameters including metabolites of glucose or lactate, electrolytes, skin temperature, and cortisol levels using microneedle patches, smart textiles, wristbands, and/or electronic epidermal tattoos[168]. Biosensors offer real-time monitoring of various functions/laboratory tests to individuals, who may control them and therefore decrease the risk of cancer-associated factors, including diabetes, hypertension, and lack of exercise.

Simple AI models are commonly more transparent, but less accurate, than complex ones. In contrast, complex models (i.e., CNNs) achieve higher accuracy but often lack interpretability. As mentioned earlier, explainable AI (XAI)[150] aims to make AI-based predictions more transparent[169,170], interpretable, and trustworthy in cancer care. XAI can reveal potential biases in AI-based predictions, strengthening their credibility. The enhanced transparency of XAI algorithms may facilitate their application in clinical decision-making and real-world clinical scenarios[171].

Discussion

The field of precision oncology may benefit greatly from the integration of AI/ML because these techniques offer a promising avenue by which to comprehend the complexity of tumor biology. Owing to the convergence of advanced AI/ML/DL data analytical techniques (software), computer hardware computational advances, high-bandwidth and cloud computing infrastructures, and innovative advanced therapeutics, we are currently at a "sea change" transition point in oncology (Fig. 2). By analyzing multidimensional -omics data, spatial pathology, and radiomics data, these technologies enable a deeper understanding of the intricate molecular pathways within tumors, aiding in the identification of critical nodes within the tumor's biology to optimize treatment selection. However, as other investigators have reported, "deployment of medical AI systems in routine clinical care presents an important yet largely unfulfilled opportunity"[172]. The applications of AI/ML in precision oncology are extensive and include the generation of synthetic data, e.g., digital twins, in order to provide the necessary information to design or expedite the conduct of clinical trials. Digital twins hold the promise of accelerating scientific discoveries and can be an important tool for decision-making based on the synergistic combination of models and data[173]. The National Academy of Sciences has defined a digital twin as a set of virtual information constructs that mimics the structure, context, and behavior of a natural, engineered, or social system (or system-of-systems), is dynamically updated with data from its physical twin, has a predictive capability, and informs decisions that realize value. The bidirectional interaction between the virtual digital twins and the physical real patients is central to the digital twin approach[173]. Currently, many operational and technical challenges remain related to data technology, engineering, and storage; algorithm development and structures; and other

Fig. 2 | Convergence of innovations in artificial intelligence analytical techniques. New modalities for deep measurement of disease and precision oncology therapeutics represent a potential "sea change" transition point for precision oncology. The first "wave" included the development of symbolic artificial intelligence tools (1997, Deep Blue expert system beat Kasparov in chess; 2011, Watson expert system won Jeopardy). These advances were followed by the second "wave", e.g., the development of deep learning tools (2012, ImageNet; 2016, AlphaGo beat Lee Sedol in GO). The third "wave" included the transformers (2018, GPT; 2020, AlphaFold2; 2022, ChatGPT, DALL-E). Simultaneously, starting in 1997, significant advances were made in biomarker innovation that enabled an improved understanding of tumor biology in parallel with accelerated drug development that involved targeted therapy and immunotherapy. "Created with BioRender.com".

elements of the data and analytical pipeline. The digitization of slides, reporting of results, and generation of new codes to submit to payers are ongoing themes that should be validated in precision oncology.

One challenge is ensuring the quality and quantity of the data. Guidelines to standardize data structure have been developed, such as the "FAIR" data principles, which stipulate that data must be findable (have adequate metadata and a persistent identifier), accessible (data and metadata are understandable to humans and machines and are deposited in a trusted repository), interoperable (metadata use a shared and broadly applicable machine language), and reusable (data have clear usage licenses, adhere to confidentiality standards, and provide accurate information on provenance)[174].

Models built on training data may not reflect the underlying heterogeneity of the patient population because they were derived from a specific patient subpopulation, non-representative of the overall population, therefore leading to biased output and limited generalizability. For example, a foundation model trained on data derived from a specific patient population may not apply to patients with different ethnic, cultural, socio-economic, and medical practice standards[175]. To ensure that highly predictive AI/ML tools are developed, there is an urgent need to share data, which is a major challenge for institutions that have traditionally considered their data as proprietary.

Another major hurdle in the ongoing application of AI/ML in practice is the seamless integration of this technology in the existing clinical workflow of patient care. Although HIPAA-compliant generative AI can be seamlessly integrated into EHRs, personalizing responses to patient messages, streamlining handoff summaries, and providing up-to-date insights for the physicians, these efficiency tools are not routinely used[176]. Furthermore, AI-enabled clinical decision support systems are too early in their development to be used. Significant work is needed to overcome barriers to AI integration associated with the cost, effort, and natural resistance to change. Data repositories are controlled by individual departments or institutions and cannot be accessed by other departments or institutions, thus hampering system interoperability[177,178]. Significant time and effort are required to educate EHR users to incorporate AI/ML technology in the clinical workflow without disruption, particularly when real-time decision-making is required[179,180]. Other barriers include the significant cost to purchase, personalize for each institution, and maintain AI/ML software. Finally, the adoption of AI-based tools should not jeopardize patient safety

or the wellbeing of hospital employees, who should not have to work overtime to ensure the smooth operation of the clinical workflow.

Significant resources are required to convert conventional pathology to a fully digital format, with significant costs associated with digitization and data storage and analysis. This transformation requires a significant campaign of education and training for its successful implementation. Additionally, there is generally an inherent impediment to the adoption of new technologies, including AI/ML. The complexity of learning a new science and applying this to the practice of medical oncology represents an enormous challenge for oncologists because computer science has not been a part of their training. Collaboration with scientists who develop AI/ML technologies and harmonization of policy and practice will be required for the integration of these new disciplines and technologies in clinical practice. Currently, formal educational programs on AI in oncology are lacking[181]. However, several courses on the application of AI are being organized to provide basic knowledge about this rapidly developing field[182,183]. Their aim is to educate clinicians to understand basic principles of AI, interpret AI-generated data, recognize the limitations of AI, and effectively utilize advanced tools for both clinical and research purposes. The development of user-friendly comprehensive training programs focusing on the integration of AI into clinical practice will be critical for the future of oncology care.

Standardization in applying AI/ML in oncology and adaptation to the new AI/ML-driven changes will be prolonged unless the current reimbursement models prioritize rapid implementation of these transformational technologies in oncology practice.

As a result of these challenges, very few clinical trials in oncology have been conducted with the prospective use of such models, although multiple articles have been published regarding the availability of this technology. Currently, the main use of AI/ML technology in precision oncology is associated with image analysis to identify radiomic features, pathologic characteristics, and other signatures/biomarkers associated with clinical outcomes. However, for the use of AI as an intelligent "agent" or medical assistant to become possible, development of relevant benchmarks to ensure performance under real-world conditions will be necessary[184].

To minimize resistance to its adoption, the transition to AI-enabled clinical practice should occur efficiently and smoothly. Strategies to reduce resistance from healthcare professionals and institutions include the involvement of all stakeholders (including physicians) in the planning, decision-making, and implementation processes; incorporation of AI in

routine insurance coverage; clinician training; and continued support to minimize concerns regarding the use of AI/ML tools. The feasibility and effectiveness of AI applications should be continuously assessed. Feedback and experience with AI tools will optimize their use in clinical practice. AI-enabled specific projects may help build trust and reduce physician burnout[185]. Finally, patient education regarding the use of AI may increase their engagement and trust in AI-tools, enhancing the application of AI in clinical practice.

The next layer of advances in AI may include the merging of the symbolic and the DL models in order to combine the benefits of both approaches, i.e., neuro-symbolic AI[186]. This will allow both the benefits of the neural networks and the available structured knowledge regarding tumor biology to be merged into a more explainable, high-performance technology.

In conclusion, considering a patient's individual characteristics and the role of multiplex, multi-omics analyses, AI-driven decision support tools will optimize treatment strategies and clinical trial enrollment, leading to better outcomes, accelerating drug development, and advancing the standard of care. However, current data can be best categorized as early "proof-of-concept" evidence. To favorably impact standard of care, these AI/ML models must go through the prospective, multicentric, large sample size demonstration of clinical validity, clinical utility, and real-world usability that is required of all new technologies, diagnostics, or therapies.

Data availability
Non-applicable.

Code availability
Non-applicable.

References

1. Bhinder, B., Gilvary, C., Madhukar, N. S. & Elemento, O. Artificial Intelligence in Cancer Research and Precision Medicine. *Cancer Discov* **11**, 900–915 (2021).
2. Somashekhar, S. P. et al. Watson for Oncology and breast cancer treatment recommendations: agreement with an expert multidisciplinary tumor board. *Ann Oncol* **29**, 418–423 (2018).
3. Jie, Z., Zhiying, Z. & Li, L. A meta-analysis of Watson for Oncology in clinical application. *Sci Rep* **11**, 5792 (2021).
4. Kris, M. G. et al. Assessing the performance of Watson for oncology, a decision support system, using actual contemporary clinical cases. *J. Clin. Oncol.* **33**, 8023 (2015).
5. Farina, E., Nabhen, J. J., Dacoregio, M. I., Batalini, F. & Moraes, F. Y. An overview of artificial intelligence in oncology. *Future Sci OA* **8**, FSO787 (2022).
6. Sarker, I. H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput Sci* **2**, 160 (2021).
7. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
8. Vaswani, A. et al. Attention Is All You Need. *ArXiv. /abs/1706.03762* (2017).
9. Bommasani, R. et al. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258* (2021).
10. Boehm, K. M., Khosravi, P., Vanguri, R., Gao, J. & Shah, S. P. Harnessing multimodal data integration to advance precision oncology. *Nat Rev Cancer* **22**, 114–126 (2022).
11. Tsimberidou, A. M. et al. T-cell receptor-based therapy: an innovative therapeutic approach for solid tumors. *J Hematol Oncol* **14**, 102 (2021).
12. Fountzilas, E., Vo, H. H., Mueller, P., Kurzrock, R. & Tsimberidou, A. M. Correlation between biomarkers and treatment outcomes in diverse cancers: a systematic review and meta-analysis of phase I and II immunotherapy clinical trials. *Eur J Cancer* **189**, 112927 (2023).
13. Hotta, T. et al. Deep learning-based diagnosis from endobronchial ultrasonography images of pulmonary lesions. *Scientific Reports* **12**, 13710 (2022).
14. Zhang, Y., Cui, J., Wan, W. & Liu, J. Multimodal imaging under artificial intelligence algorithm for the diagnosis of liver cancer and its relationship with expressions of EZH2 and p57. *Computational Intelligence Neuroscience* **2022**, 4081654 (2022).
15. Ren, C. et al. Clinico-biological-radiomics (CBR) based machine learning for improving the diagnostic accuracy of FDG-PET false-positive lymph nodes in lung cancer. *European Journal Medical Research* **28**, 554 (2023).
16. Luo, X. et al. Automated segmentation of brain metastases with deep learning: A multi-center, randomized crossover, multi-reader evaluation study. *Neuro-oncology* **26**, 2140–2151 (2024).
17. Le, Y. et al. CT radiomics analysis discriminates pulmonary lesions in patients with pulmonary MALT lymphoma and non-pulmonary MALT lymphoma. *Methods* **224**, 54–62 (2024).
18. Arezzo, F. et al. A machine learning approach applied to gynecological ultrasound to predict progression-free survival in ovarian cancer patients. *Archives Gynecology Obstetrics* **306**, 2143–2154 (2022).
19. Zhang, K. et al. Using deep learning to predict survival outcome in non-surgical cervical cancer patients based on pathological images. *Journal Cancer Research Clinical Oncology* **149**, 6075–6083 (2023).
20. Jiang, W. et al. A nomogram based on a collagen feature support vector machine for predicting the treatment response to neoadjuvant chemoradiotherapy in rectal cancer patients. *Annals Surgical Oncology* **28**, 6408–6421 (2021).
21. Sharma, A. et al. Development and prognostic validation of a three-level NHG-like deep learning-based model for histological grading of breast cancer. *Breast Cancer Research* **26**, 17 (2024).
22. Arabyarmohammadi, S. et al. Machine learning to predict risk of relapse using cytologic image markers in patients with acute myeloid leukemia posthematopoietic cell transplantation. *JCO Clinical Cancer Informatics* **6**, e2100156 (2022).
23. Sundar, R. et al. Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomised phase III SAMIT trial. *Gut* **71**, 676–685 (2022).
24. Christopoulos, P. et al. Plasma Proteome–Based Test for First-Line Treatment Selection in Metastatic Non–Small Cell Lung Cancer. *JCO Precision Oncology* **8**, e2300555 (2024).
25. Malhaire, C. et al. Exploring the added value of pretherapeutic MR descriptors in predicting breast cancer pathologic complete response to neoadjuvant chemotherapy. *European Radiology* **33**, 8142–8154 (2023).
26. Lv, L. et al. Radiomic analysis for predicting prognosis of colorectal cancer from preoperative 18F-FDG PET/CT. *Journal translational medicine* **20**, 66 (2022).
27. Kim, C. G. et al. A phase II open-label randomized clinical trial of preoperative durvalumab or durvalumab plus tremelimumab in resectable head and neck squamous cell carcinoma. *Clinical Cancer Research* **30**, 2097–2110 (2024).
28. Sobottka, B. et al. Establishing standardized immune phenotyping of metastatic melanoma by digital pathology. *Laboratory investigation* **101**, 1561–1570 (2021).
29. Manz, C. R. et al. Effect of integrating machine learning mortality estimates with behavioral nudges to clinicians on serious illness conversations among patients with cancer: a stepped-wedge cluster randomized clinical trial. *JAMA oncology* **6**, e204759–e204759 (2020).
30. Guével, E. et al. Développement d'un modèle de traitement automatique du langage pour calculer des indicateurs qualité du cancer du sein: une étude transversale multicentrique. *Revue D'epidemiologie et de Sante Publique* **71**, 102189–102189 (2023).

31. Ma, L. et al. Correlation between AI-based CT organ features and normal lung dose in adjuvant radiotherapy following breast-conserving surgery: a multicenter prospective study. *BMC cancer* **23**, 1085 (2023).

32. Chrystall, D. et al. Deep learning enables MV-based real-time image guided radiation therapy for prostate cancer patients. *Physics Medicine Biology* **68**, 095016 (2023).

33. Huang, Z. et al. Artificial intelligence reveals features associated with breast cancer neoadjuvant chemotherapy responses from multi-stain histopathologic images. *NPJ Precis Oncol* **7**, 14 (2023).

34. Dekker, T. J. et al. Determining sensitivity and specificity of HER2 testing in breast cancer using a tissue micro-array approach. *Breast Cancer Res* **14**, R93 (2012).

35. Koomen, B. M. et al. False-negative programmed death-ligand 1 immunostaining in ethanol-fixed endobronchial ultrasound-guided transbronchial needle aspiration specimens of non-small-cell lung cancer patients. *Histopathology* **79**, 480–490 (2021).

36. Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V. & Madabhushi, A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* **16**, 703–715 (2019).

37. Jiang, Y., Yang, M., Wang, S., Li, X. & Sun, Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun (Lond)* **40**, 154–166 (2020).

38. Baxi, V. et al. Association of artificial intelligence-powered and manual quantification of programmed death-ligand 1 (PD-L1) expression with outcomes in patients treated with nivolumab +/- ipilimumab. *Mod Pathol* **35**, 1529–1539 (2022).

39. Chatrian, A. et al. Artificial intelligence for advance requesting of immunohistochemistry in diagnostically uncertain prostate biopsies. *Mod Pathol* **34**, 1780–1794 (2021).

40. Cheng, G. et al. Artificial Intelligence-Assisted Score Analysis for Predicting the Expression of the Immunotherapy Biomarker PD-L1 in Lung Cancer. *Front Immunol* **13**, 893198 (2022).

41. Wu, J. et al. Artificial intelligence-assisted system for precision diagnosis of PD-L1 expression in non-small cell lung cancer. *Mod Pathol* **35**, 403–411 (2022).

42. Abele, N. et al. Noninferiority of Artificial Intelligence-Assisted Analysis of Ki-67 and Estrogen/Progesterone Receptor in Breast Cancer Routine Diagnostics. *Mod Pathol* **36**, 100033 (2023).

43. Mercan, C. et al. Deep learning for fully-automated nuclear pleomorphism scoring in breast cancer. *NPJ Breast Cancer* **8**, 120 (2022).

44. Widmaier, M. et al. Comparison of continuous measures across diagnostic PD-L1 assays in non-small cell lung cancer using automated image analysis. *Modern Pathology* **33**, 380–390 (2020).

45. Wang, X. et al. How can artificial intelligence models assist PD-L1 expression scoring in breast cancer: results of multi-institutional ring studies. *NPJ Breast Cancer* **7**, 61 (2021).

46. Fourkioti, O., De Vries, M. & Bakal, C. CAMIL: Context-Aware Multiple Instance Learning for Cancer Detection and Subtyping in Whole Slide Images. *arXiv preprint arXiv:2305.05314* (2023).

47. Wang, X. et al. Dual-scale categorization based deep learning to evaluate programmed cell death ligand 1 expression in non-small cell lung cancer. *Medicine (Baltimore)* **100**, e25994 (2021).

48. Chang, S., Park, H. K., Choi, Y. L. & Jang, S. J. Interobserver Reproducibility of PD-L1 Biomarker in Non-small Cell Lung Cancer: A Multi-Institutional Study by 27 Pathologists. *J Pathol Transl Med.* **53**, 347–353 (2019).

49. Vandenberghe, M. E. et al. Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. *Sci Rep* **7**, 45938 (2017).

50. Wang, X. et al. Prediction of BRCA Gene Mutation in Breast Cancer Based on Deep Learning and Histopathology Images. *Front Genet* **12**, 661109 (2021).

51. Jang, H. J., Lee, A., Kang, J., Song, I. H. & Lee, S. H. Prediction of clinically actionable genetic alterations from colorectal cancer histopathology images using deep learning. *World J Gastroenterol* **26**, 6207–6223 (2020).

52. Liu, H. et al. Breast Cancer Molecular Subtype Prediction on Pathological Images with Discriminative Patch Selection and Multi-Instance Learning. *Front Oncol* **12**, 858453 (2022).

53. Fu, Y. et al. Deep learning predicts patients outcome and mutations from digitized histology slides in gastrointestinal stromal tumor. *NPJ Precis Oncol* **7**, 71 (2023).

54. Diao, J. A. et al. Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes. *Nat Commun* **12**, 1613 (2021).

55. Kather, J. N. et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* **25**, 1054–1056 (2019).

56. Echle, A. et al. Clinical-Grade Detection of Microsatellite Instability in Colorectal Tumors by Deep Learning. *Gastroenterology* **159**, 1406–1416.e1411 (2020).

57. Yamashita, R. et al. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol* **22**, 132–141 (2021).

58. Chen, M. et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. *NPJ Precis Oncol* **4**, 14 (2020).

59. Kather, J. N. et al. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* **1**, 789–799 (2020).

60. Saldanha, O. L. et al. Self-supervised attention-based deep learning for pan-cancer mutation prediction from histopathology. *NPJ Precis Oncol* **7**, 35 (2023).

61. Fremond, S. et al. Interpretable deep learning model to predict the molecular classification of endometrial cancer from haematoxylin and eosin-stained whole-slide images: a combined analysis of the PORTEC randomised trials and clinical cohorts. *Lancet Digit Health* **5**, e71–e82 (2023).

62. Tsai, P. C. et al. Histopathology images predict multi-omics aberrations and prognoses in colorectal cancer patients. *Nat Commun* **14**, 2102 (2023).

63. Esteva, A. et al. Prostate cancer therapy personalization via multi-modal deep learning on randomized phase III clinical trials. *NPJ Digit Med.* **5**, 71 (2022).

64. Mobadersany, P. et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci USA* **115**, E2970–E2979 (2018).

65. Zeng, H., Chen, L., Zhang, M., Luo, Y. & Ma, X. Integration of histopathological images and multi-dimensional omics analyses predicts molecular features and prognosis in high-grade serous ovarian cancer. *Gynecol Oncol* **163**, 171–180 (2021).

66. Azadi, P. et al. ALL-IN: AL ocal GL obal Graph-Based DI stillatio N Model for Representation Learning of Gigapixel Histopathology Images With Application In Cancer Risk Assessment. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 765–775 (Springer, 2023).

67. Park, S. et al. Artificial Intelligence-Powered Spatial Analysis of Tumor-Infiltrating Lymphocytes as Complementary Biomarker for Immune Checkpoint Inhibition in Non-Small-Cell Lung Cancer. *J Clin Oncol* **40**, 1916–1928 (2022).

68. Zimmermann, E. et al. Virchow 2: Scaling Self-Supervised Mixed Magnification Models in Pathology. *arXiv preprint arXiv:2408.00738* (2024).

69. Clark, K. et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *Journal digital imaging* **26**, 1045–1057 (2013).

70. Poirion, O. B., Jing, Z., Chaudhary, K., Huang, S. & Garmire, L. X. DeepProg: an ensemble of deep-learning and machine-learning

models for prognosis prediction using multi-omics data. *Genome Med* **13**, 112 (2021).

71. Zhu, B. et al. Integrating Clinical and Multiple Omics Data for Prognostic Assessment across Human Cancers. *Sci Rep* **7**, 16954 (2017).

72. Lobato-Delgado, B., Priego-Torres, B. & Sanchez-Morillo, D. Combining Molecular, Imaging, and Clinical Data Analysis for Predicting Cancer Prognosis. *Cancers (Basel)* **14**, 3215 (2022).

73. Neely, B. A. et al. Toward an Integrated Machine Learning Model of a Proteomics Experiment. *J Proteome Res.* **22**, 681–696 (2023).

74. Cheng, M. W., Mitra, M. & Coller, H. A. Pan-cancer landscape of epigenetic factor expression predicts tumor outcome. *Commun Biol* **6**, 1138 (2023).

75. Bahado-Singh, R. O. et al. Precision gynecologic oncology: circulating cell free DNA epigenomic analysis, artificial intelligence and the accurate detection of ovarian cancer. *Sci Rep* **12**, 18625 (2022).

76. Albaradei, S. et al. MetaCancer: A deep learning-based pan-cancer metastasis prediction model developed using multi-omics data. *Comput Struct Biotechnol J* **19**, 4404–4411 (2021).

77. Godlewski, A. et al. A comparison of different machine-learning techniques for the selection of a panel of metabolites allowing early detection of brain tumors. *Sci Rep* **13**, 11044 (2023).

78. Bifarin, O. O. et al. Machine Learning-Enabled Renal Cell Carcinoma Status Prediction Using Multiplatform Urine-Based Metabolomics. *J Proteome Res.* **20**, 3629–3641 (2021).

79. Sorin, M. et al. Single-cell spatial landscapes of the lung tumour immune microenvironment. *Nature* **614**, 548–554 (2023).

80. Karimi, E. et al. Single-cell spatial immune landscapes of primary and metastatic brain tumours. *Nature* **614**, 555–563 (2023).

81. Keren, L. et al. A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373–1387.e1319 (2018).

82. Jimenez-Sanchez, D. et al. Weakly supervised deep learning to predict recurrence in low-grade endometrial cancer from multiplexed immunofluorescence images. *NPJ Digit Med.* **6**, 48 (2023).

83. Mund, A. et al. Deep Visual Proteomics defines single-cell identity and heterogeneity. *Nat Biotechnol* **40**, 1231–1240 (2022).

84. Song, H. et al. Survival stratification for colorectal cancer via multi-omics integration using an autoencoder-based model. *Exp Biol Med (Maywood)* **247**, 898–909 (2022).

85. Battistella, E. et al. COMBING: Clustering in Oncology for Mathematical and Biological Identification of Novel Gene Signatures. *IEEE/ACM Trans Comput Biol Bioinform* **19**, 3317–3331 (2022).

86. Nakajima, E. C. et al. Tumor Size Is Not Everything: Advancing Radiomics as a Precision Medicine Biomarker in Oncology Drug Development and Clinical Care. A Report of a Multidisciplinary Workshop Coordinated by the RECIST Working Group. *JCO Precision Oncology* **8**, e2300687 (2024).

87. Cobo, M., Menéndez Fernández-Miranda, P., Bastarrika, G. & Lloret Iglesias, L. Enhancing radiomics and Deep Learning systems through the standardization of medical imaging workflows. *Scientific data* **10**, 732 (2023).

88. Lerousseau, M. et al. Weakly supervised pan-cancer segmentation tool. in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24* 248-256 (Springer, 2021).

89. Christ, P. F. et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19* 415-423 (Springer, 2016).

90. Kamnitsas, K. et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical image analysis* **36**, 61–78 (2017).

91. Roth, H. R. et al. Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part I 18* 556-564 (Springer, 2015).

92. Mata, L. A., Retamero, J. A., Gupta, R. T., García Figueras, R. & Luna, A. Artificial intelligence–assisted prostate cancer diagnosis: Radiologic-pathologic correlation. *Radiographics* **41**, 1676–1697 (2021).

93. Turkbey, B. & Haider, M. A. Artificial intelligence for automated cancer detection on prostate MRI: opportunities and ongoing challenges, from the AJR special series on AI applications. *American Journal Roentgenology* **219**, 188–194 (2022).

94. Hurkmans, C. et al. A joint ESTRO and AAPM guideline for development, clinical validation and reporting of artificial intelligence models in radiation therapy. *Radiother Oncol* **197**, 110345 (2024).

95. Park, C. et al. Tumor-infiltrating lymphocyte enrichment predicted by CT radiomics analysis is associated with clinical outcomes of non-small cell lung cancer patients receiving immune checkpoint inhibitors. *Front Immunol* **13**, 1038089 (2022).

96. Roisman, L. C. et al. Radiological artificial intelligence - predicting personalized immunotherapy outcomes in lung cancer. *NPJ Precis Oncol* **7**, 125 (2023).

97. Qi, L. et al. CT radiomics-based biomarkers can predict response to immunotherapy in hepatocellular carcinoma. *Sci Rep* **14**, 20027 (2024).

98. Kazmierski, M. et al. Multi-institutional Prognostic Modeling in Head and Neck Cancer: Evaluating Impact and Generalizability of Deep Learning and Radiomics. *Cancer Res Commun* **3**, 1140–1151 (2023).

99. Peng, J. et al. A novel sub-regional radiomics model to predict immunotherapy response in non-small cell lung carcinoma. *J Transl Med* **22**, 87 (2024).

100. Dercle, L. et al. Early Readout on Overall Survival of Patients With Melanoma Treated With Immunotherapy Using a Novel Imaging Analysis. *JAMA Oncol* **8**, 385–392 (2022).

101. Wang, J. et al. CT radiomics-based model for predicting TMB and immunotherapy response in non-small cell lung cancer. *BMC Med Imaging* **24**, 45 (2024).

102. Zhao, J. et al. Radiomic and clinical data integration using machine learning predict the efficacy of anti-PD-1 antibodies-based combinational treatment in advanced breast cancer: a multicentered study. *Journal ImmunoTherapy Cancer* **11**, e006514 (2023).

103. Lou, B. et al. An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction. *Lancet Digital Health* **1**, e136–e147 (2019).

104. Kann, B. H. et al. Screening for extranodal extension in HPV-associated oropharyngeal carcinoma: evaluation of a CT-based deep learning algorithm in patient data from a multicentre, randomised de-escalation trial. *Lancet Digital Health* **5**, e360–e369 (2023).

105. Xu X. et al. Workshop on Machine Learning for Systems. (2023).

106. Rußwurm, M. et al. Machine Learning for Remote Sensing (ML4RS). in *ICLR 2024 Workshops*.

107. Brendlin, A. S. et al. A Machine learning model trained on dual-energy CT radiomics significantly improves immunotherapy response prediction for patients with stage IV melanoma. *J. Immunother. Cancer* **9**, e003261 (2021).

108. Hu, W. et al. A comparison study of artificial intelligence performance against physicians in benign-malignant classification of pulmonary nodules. *Oncologie* **26**, 581–586 (2024).

109. Kim, H. E. et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *Lancet Digit Health* **2**, e138–e148 (2020).

110. Mota, S. M. et al. Artificial Intelligence Improves the Ability of Physicians to Identify Prostate Cancer Extent. *J Urol* **212**, 52–62 (2024).

111. Klontzas, M. E. Radiomics feature reproducibility: The elephant in the room. *Eur J Radiol* **175**, 111430 (2024).

112. Gomes, B. & Ashley, E. A. Artificial Intelligence in Molecular Medicine. *N. Engl J Med.* **388**, 2456–2465 (2023).

113. Poplin, R. et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnol* **36**, 983–987 (2018).

114. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491–498 (2011).

115. Rauschert, S., Raubenheimer, K., Melton, P. E. & Huang, R. C. Machine learning and clinical epigenetics: a review of challenges for diagnosis and classification. *Clin Epigenetics* **12**, 51 (2020).

116. Hao, X. et al. DNA methylation markers for diagnosis and prognosis of common cancers. *Proc Natl Acad Sci USA* **114**, 7414–7419 (2017).

117. Qiu, J. et al. CpG Methylation Signature Predicts Recurrence in Early-Stage Hepatocellular Carcinoma: Results From a Multicenter Study. *J Clin Oncol* **35**, 734–742 (2017).

118. Brasil, S. et al. Artificial Intelligence in Epigenetic Studies: Shedding Light on Rare Diseases. *Front Mol Biosci* **8**, 648012 (2021).

119. Keyl, P. et al. Patient-level proteomic network prediction by explainable artificial intelligence. *NPJ Precis Oncol* **6**, 35 (2022).

120. Rashno, E., Eskandari, A., Anand, A. & Zulkernine, F. Survey: Transformer-based Models in Data Modality Conversion. *arXiv preprint arXiv:2408.04723* (2024).

121. Lipkova, J. et al. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* **40**, 1095–1110 (2022).

122. Gao, Y. & Alison Noble, J. Detection and Characterization of the Fetal Heartbeat in Free-hand Ultrasound Sweeps with Weakly-supervised Two-streams Convolutional Networks. 305-313 (Springer International Publishing, Cham, 2017).

123. Nie, D. et al. Medical Image Synthesis with Context-Aware Generative Adversarial Networks. 417–425 (Springer International Publishing, Cham, 2017).

124. Kim, H. et al. Clinical feasibility of deep learning-based synthetic CT images from T2-weighted MR images for cervical cancer patients compared to MRCAT. *Sci Rep* **14**, 8504 (2024).

125. Barber, R. D. & Kroeger, K. Towards Network Medicine: Implementation of Panomics and Artificial Intelligence for Precision Medicine. in *Digital Disruption in Healthcare* 27–43 (Springer, 2022).

126. Acosta, J. N., Falcone, G. J., Rajpurkar, P. & Topol, E. J. Multimodal biomedical AI. *Nat Med.* **28**, 1773–1784 (2022).

127. Vanguri, R. S. et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer* **3**, 1151–1164 (2022).

128. Zhang, Y., Xu, Y., Chen, J., Xie, F. & Chen, H. Prototypical information bottlenecking and disentangling for multimodal cancer survival prediction. *arXiv preprint arXiv:2401.01646* (2024).

129. Lederman, A., Lederman, R. & Verspoor, K. Tasks as needs: reframing the paradigm of clinical natural language processing research for real-world decision support. *J Am Med Inform Assoc* **29**, 1810–1817 (2022).

130. Webster, P. Six ways large language models are changing healthcare. *Nat Med.* **29**, 2969–2971 (2023).

131. Benary, M. et al. Leveraging Large Language Models for Decision Support in Personalized Oncology. *JAMA Netw Open* **6**, e2343689 (2023).

132. Moazzam, Z., Cloyd, J., Lima, H. A. & Pawlik, T. M. Quality of ChatGPT Responses to Questions Related to Pancreatic Cancer and its Surgical Care. *Ann Surg Oncol* **30**, 6284–6286 (2023).

133. Sorin, V. et al. Large language model (ChatGPT) as a support tool for breast tumor board. *NPJ Breast Cancer* **9**, 44 (2023).

134. Zakka, C. et al. Almanac - retrieval-augmented language models for clinical medicine. *NEJM AI* **1**, Aloa2300068 (2024).

135. Singhal, K. et al. Towards expert-level medical question answering with large language models. *Nature Medicine* 1–8 (2025).

136. Tu, T. et al. Towards Generalist Biomedical AI. *NEJM AI* **1**, Aloa2300138 (2024).

137. Yang, X. et al. A large language model for electronic health records. *NPJ Digit Med.* **5**, 194 (2022).

138. Rajkomar, A. et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med.* **1**, 18 (2018).

139. Chen, S. et al. Use of Artificial Intelligence Chatbots for Cancer Treatment Information. *JAMA Oncol* **9**, 1459–1462 (2023).

140. Lee, P., Bubeck, S. & Petro, J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *N. Engl J Med.* **388**, 1233–1239 (2023).

141. U.S. Food & Drugs Administration. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. Vol. 2025 (2025).

142. Laishram, I. S. Data bias in precision medicine. *International Journal Advances Medicine* **9**, 1239 (2022).

143. Kim, J., Cai, Z. R., Chen, M. L., Simard, J. F. & Linos, E. Assessing Biases in Medical Decisions via Clinician and AI Chatbot Responses to Patient Vignettes. *JAMA Netw Open* **6**, e2338050 (2023).

144. Ning, Y. et al. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. *Lancet Digit Health* **6**, e848–e856 (2024).

145. Murphy, K. et al. Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Med Ethics* **22**, 14 (2021).

146. Petersson, L. et al. Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. *BMC Health Serv Res.* **22**, 850 (2022).

147. Farasati Far, B. Artificial intelligence ethics in precision oncology: balancing advancements in technology with patient privacy and autonomy. *Explor Target Antitumor Ther* **4**, 685–689 (2023).

148. Shreve, J. T., Khanani, S. A. & Haddad, T. C. Artificial Intelligence in Oncology: Current Capabilities, Future Opportunities, and Ethical Considerations. *Am Soc Clin Oncol Educ Book* **42**, 1–10 (2022).

149. Sheu, R. K. & Pardeshi, M. S. A Survey on Medical Explainable AI (XAI): Recent Progress, Explainability Approach, Human Interaction and Scoring System. *Sensors (Basel)* **22**, 8068 (2022).

150. Klauschen, F. et al. Toward Explainable Artificial Intelligence for Precision Pathology. *Annu Rev Pathol* **19**, 541–570 (2024).

151. Zhang, J. & Zhang, Z. M. Ethics and governance of trustworthy medical artificial intelligence. *BMC Med Inform Decis Mak* **23**, 7 (2023).

152. Reddy, S. et al. Evaluation framework to guide implementation of AI systems into healthcare settings. *BMJ Health Care Inform* **28**, e100444 (2021).

153. Obafemi-Ajayi, T. et al. No-boundary thinking: a viable solution to ethical data-driven AI in precision medicine. *AI Ethics* **2**, 635–643 (2022).

154. Loftus, T. J. et al. Federated learning for preserving data privacy in collaborative healthcare research. *Digital Health* **8**, 20552076221134455 (2022).

155. Rinaldi, E. et al. International clinical research data ecosystem: from data standardization to federated analysis. in *Telehealth Ecosystems in Practice* 133–134 (IOS Press, 2023).

156. Sheller, M. J. et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific reports* **10**, 12598 (2020).

157. Casaletto, J., Bernier, A., McDougall, R. & Cline, M. S. Federated Analysis for Privacy-Preserving Data Sharing: A Technical and Legal Primer. *Annual review genomics human genetics* **24**, 347–368 (2023).

158. Health, N.I.o. Fast Healthcare Interoperability Resources (FHIR) Standard. *NIH guide to grants and contracts. NOT-OD-19-122.[Internet]* (2019).

159. Linh, V. T. N. et al. 3D plasmonic hexaplex paper sensor for label-free human saliva sensing and machine learning-assisted early-stage lung cancer screening. *Biosensors Bioelectronics* **244**, 115779 (2024).

160. Elsheakh, D. N., Mohamed, R. A., Fahmy, O. M., Ezzat, K. & Eldamak, A. R. Complete breast cancer detection and monitoring system by using microwave textile based antenna sensors. *Biosensors* **13**, 87 (2023).

161. Park, S., Park, H. J. & Lee, K. H. Multi-Marker Biosensor Integrated Explainable AI-Based Cancer Screening System. in *Electrochemical Society Meeting Abstracts prime2024* 4995-4995 (The Electrochemical Society, Inc., 2024).

162. Jo, K. et al. Machine learning-assisted label-free colorectal cancer diagnosis using plasmonic needle-endoscopy system. *Biosensors Bioelectronics* **264**, 116633 (2024).

163. Hu, J. et al. Smell cancer by machine learning-assisted peptide/MXene bioelectronic array. *Biosensors Bioelectronics* **262**, 116562 (2024).

164. Lei, Y. et al. Detection of carcinoembryonic antigen specificity using microwave biosensor with machine learning. *Biosensors Bioelectronics* **269**, 116908 (2025).

165. Zhang, C. et al. Machine learning assisted dual-modal SERS detection for circulating tumor cells. *Biosensors Bioelectronics* **268**, 116897 (2025).

166. Zhang, J. et al. Molecular separation-assisted label-free SERS combined with machine learning for nasopharyngeal cancer screening and radiotherapy resistance prediction. *Journal Photochemistry Photobiology B: Biology* **257**, 112968 (2024).

167. Asare-Werehene, M. et al. The application of an extracellular vesicle-based biosensor in early diagnosis and prediction of chemoresponsiveness in ovarian cancer. *Cancers* **15**, 2566 (2023).

168. Mishra, S. Electroceuticals in medicine–The brave new future. *Indian heart journal* **69**, 685–686 (2017).

169. Patidar, N. et al. Transparency in AI Decision Making: A Survey of Explainable AI Methods and Applications. *Adv. Robot. Technol.* **2**, 000110 (2024).

170. Tiwari, R. Explainable ai (xai) and its applications in building trust and understanding in AI decision making. *International J. Sci. Res. Eng. Manag* **7**, 1–13 (2023).

171. Edwards, L. & Veale, M. Enslaving the algorithm: From a "right to an explanation" to a "right to better decisions"? *IEEE Security Privacy* **16**, 46–54 (2018).

172. Rajpurkar, P., Chen, E., Banerjee, O. & Topol, E. J. AI in health and medicine. *Nat Med.* **28**, 31–38 (2022).

173. Cooke, P. Image and reality:'digital twins' in smart factory automotive process innovation–critical issues. *Regional Studies* **55**, 1630–1641 (2021).

174. Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016).

175. Mittermaier, M., Raza, M. M. & Kvedar, J. C. Bias in AI-based models for medical applications: challenges and mitigation strategies. *NPJ Digit Med* **6**, 113 (2023).

176. EPIC. Artificial Intelligence: https://www.epic.com/software/ai/.

177. Szarfman, A. et al. Recommendations for achieving interoperable and shareable medical data in the USA. *Communications Medicine* **2**, 86 (2022).

178. Calvino, G. et al. Federated Learning: Breaking Down Barriers in Global Genomic Research. *Genes* **15**, 1650 (2024).

179. Kawamoto, K., Finkelstein, J. & Del Fiol, G. Implementing machine learning in the electronic health record: checklist of essential considerations. in *Mayo Clinic Proceedings*, Vol. 98 366-369 (Elsevier, 2023).

180. Rose, C. & Chen, J. H. Learning from the EHR to implement AI in healthcare. *npj Digital Medicine* **7**, 330 (2024).

181. Prelaj, A., Scoazec, G., Ferber, D. & Kather, J. Oncology education in the age of artificial intelligence. *ESMO Real World Data Digital Oncology* **6**, 100079 (2024).

182. Kang, J. et al. National cancer institute workshop on artificial intelligence in radiation oncology: training the next generation. *Practical radiation oncology* **11**, 74–83 (2021).

183. Toussaint, J.-M. S. A.-A. New training courses on artificial intelligence offered to healthcare professionals working in oncology. (The Canadian Cancer Society (CCS), March 30, 2022).

184. Mehandru, N. et al. Evaluating large language models as agents in the clinic. *npj Digital Medicine* **7**, 84 (2024).

185. Shuaib, A. Transforming Healthcare with AI: Promises, Pitfalls, and Pathways Forward. *Int. J. Gen. Med.* **17**, 1765–1771 (2024).

186. Sheth, A. & Roy, K. Neurosymbolic Value-Inspired Artificial Intelligence (Why, What, and How). *IEEE Intelligent Systems* **39**, 5–11 (2024).

187. Bzdok, D., Engemann, D. & Thirion, B. Inference and prediction diverge in biomedicine. *Patterns* **1**, 100119 (2020).

188. Breiman, L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science* **16**, 199–231 (2001).

189. Luo, W. et al. Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View. *J Med Internet Res.* **18**, e323 (2016).

190. Esteva, A. et al. A guide to deep learning in healthcare. *Nat Med.* **25**, 24–29 (2019).

191. Liu, Y., et al. Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv:1703.02442* (2017).

192. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *nature* **542**, 115–118 (2017).

193. Lay, N. et al. Detection of prostate cancer in multiparametric MRI using random forest with instance weighting. *Journal Medical Imaging* **4**, 024506–024506 (2017).

194. Zhang, C. et al. Toward an expert level of lung cancer detection and classification using a deep convolutional neural network. *oncologist* **24**, 1159–1165 (2019).

195. McKinney, S. M. et al. International evaluation of an AI system for breast cancer screening. *Nature* **577**, 89–94 (2020).

196. Savage, N. Tapping into the drug discovery potential of AI. *Nature. com*. https://www.nature.com/articles/d43747-021-00045-7 (2021).

197. Zhu, S., Gilbert, M., Chetty, I. & Siddiqui, F. The 2021 landscape of FDA-approved artificial intelligence/machine learning-enabled medical devices: An analysis of the characteristics and intended use. *International journal medical informatics* **165**, 104828 (2022).

198. U.S. Food and Drug Administration. FDA Authorizes Marketing of First Device that Uses Artificial Intelligence to Help Detect Potential Signs of Colon Cancer. (2021).

199. U.S. Food and Drug Administration. Optellum Virtual Nodule Clinic, Optellum software, Optellum platform. (2021).

200. U.S. Food and Drug Administration. Evaluation of Automatic Class III Designation for Paige Prostate. (2021).

201. Ye, M. et al. A classifier for improving early lung cancer diagnosis incorporating artificial intelligence and liquid biopsy. *Frontiers oncology* **12**, 853801 (2022).

202. U.S. Food and Drug Administration. ProstatID: Radiological Computer Assisted Detection/Diagnosis Software For Lesions Suspicious For Cancer (2022).

203. National Institutes of Health. AI-Driven Software Based Off NIH Algorithm Receives FDA Clearance in Detection and Diagnosis of Prostate Cancer. (2022).

204. Liu, M. et al. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. *PLoS One* **18**, e0273445 (2023).

205. Spratt, D. E. et al. Artificial intelligence predictive model for hormone therapy use in prostate cancer. *NEJM evidence* **2**, EVIDoa2300023 (2023).

206. U.S. Food and Drug Administration. Thirona BV: LungQ v3.0.0. Vol. 2025 (2024).

207. U.S. Food & Drug Administration. Avenda Health AI Prostate Cancer Planning Software. (2022).

208. Guardant Health. Shield For Detection of Colorectal Cancer Sponsor Executive Summary Molecular And Clinical Genetics Panel. Vol. 2024 (2024).
209. Schaeffer, E. M. et al. NCCN Guidelines® Insights: Prostate Cancer, Version 3.2024: Featured Updates to the NCCN Guidelines. *Journal National Comprehensive Cancer Network* **22**, 140–150 (2024).
210. Witkowski, A. M. et al. Clinical Utility of a Digital Dermoscopy Image-Based Artificial Intelligence Device in the Diagnosis and Management of Skin Cancer by Dermatologists. *Cancers* **16**, 3592 (2024).
211. U.S. Food and Drug Administration. Software-Aided Adjunctive Diagnostic Device For Use By Physicians On Lesions Suspicious For Skin Cancer. (2024).

Acknowledgements

Author contributions

A.M.T. conceived the presented paper, provided funding, and supervised the work; E.F., T.P., M.A.B., and A.C. wrote the main manuscript text; and M.A.B., T.P., and A.C. prepared the Figures. All authors have read and approved the manuscript.

Competing interests

A.M.T. declares receipt of Clinical Trial Research Funding (received through the institution): OBI Pharma, Agenus, Vividion, Macrogenics, AbbVie, IMMATICS, Novocure, Tachyon, Parker Institute for Cancer Immunotherapy, Tempus, and Tvardi; fees for consulting or advisory roles for Avstera Therapeutics, Bioeclipse, BrYet, Diaccurate, Macrogenics, NEX-I, and VinceRx. E.F. declares advisory role of Amgen LEO Pharma; travel grants from Merck, Pfizer, AstraZeneca, DEMO and K.A.M. Oncology/Hematology; Speaker fees from Roche, Leo, Pfizer, AstraZeneca, Amgen; and Stock ownership from Genprex Inc., Deciphera Pharmaceuticals, Inc. The remaining authors declare no relevant conflict of interest.

Additional information

Correspondence and requests for materials should be addressed to Apostolia M. Tsimberidou.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.